

UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE CIÊNCIA DA INFORMAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM GESTÃO & ORGANIZAÇÃO DO
CONHECIMENTO

EDUARDO CARDOSO MELO

**ANÁLISE ONTOLÓGICA DO PROCESSO DE CONCESSÃO DE AUXÍLIO
SOCIOECONÔMICO, EVASÃO E DESEMPENHO ACADÊMICO NO ENSINO
SUPERIOR DO IFMG**

Belo Horizonte

2025

EDUARDO CARDOSO MELO

**ANÁLISE ONTOLÓGICA DO PROCESSO DE CONCESSÃO DE AUXÍLIO
SOCIOECONÔMICO, EVASÃO E DESEMPENHO ACADÊMICO NO ENSINO
SUPERIOR DO IFMG**

Tese apresentada ao Programa de Pós-Graduação em Gestão & Organização do Conhecimento, Escola de Ciência da Informação da Universidade Federal de Minas Gerais, como requisito para obtenção do grau de Doutor em Gestão e Organização do Conhecimento.

Área de concentração: Ciência da Informação.

Linha de Pesquisa: Gestão e Tecnologia

Orientador: Dr. Marcello Peixoto Bax

Belo Horizonte

2025

M528a

Melo, Eduardo Cardoso.

Análise ontológica do processo de concessão de auxílio socioeconômico, evasão e desempenho acadêmico no ensino superior do IFMG [recurso eletrônico] / Eduardo Cardoso Melo. - 2025.

1 recurso eletrônico (273 f. : il., color.) : pdf.

Orientador: Marcello Peixoto Bax.

Tese (doutorado)– Universidade Federal de Minas Gerais, Escola de Ciência da Informação.

Referências: f. 228-263.

Apêndice: f. 264-273.

Exigência do sistema: Adobe Acrobat Reader.

1. Ciência da informação – Teses. 2. Estudantes – Auxílio – Teses. 3. Evasão escolar – Teses. 4. Ensino superior – Teses. 5. Ontologias (Recuperação da informação) – Teses. I. Bax, Marcello Peixoto. II. Universidade Federal de Minas Gerais. Escola de Ciência da Informação. III. Título.

CDU: 02:371.2

Ficha catalográfica: Elaine Diamantino Oliveira - CRB: 6/2742

Biblioteca Profª Etelvina Lima, Escola de Ciência da Informação da UFMG



UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE CIÊNCIA DA INFORMAÇÃO - ECI
PROGRAMA DE PÓS-GRADUAÇÃO EM GESTÃO E ORGANIZAÇÃO DO CONHECIMENTO - PPGGOC

FOLHA DE APROVAÇÃO

ORIENTAÇÃO A DADOS E ONTOLOGIAS: ANÁLISE DAS RELAÇÕES ENTRE A CONCESSÃO DE AUXÍLIOS SOCIOECONÔMICOS, EVASÃO ESCOLAR E DESEMPENHO ACADÊMICO NO IFMG

EDUARDO CARDOSO MELO

Tese submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em GESTÃO E ORGANIZAÇÃO DO CONHECIMENTO, como requisito para obtenção do grau de Doutor em GESTÃO E ORGANIZAÇÃO DO CONHECIMENTO, área de concentração CIÊNCIA DA INFORMAÇÃO, linha de pesquisa Gestão e Tecnologia da Informação e Comunicação.

Aprovada em 18 de março de 2025, por videoconferência, pela banca constituída pelos membros:

Prof(a). Marcello Peixoto Bax (Orientador)
ECI/UFMG

Dr(a). Ana Kelly Arantes
IFMG

Prof(a). José Eduardo Santarem Segundo
USP

Prof(a). José Maria Parente de Oliveira
ITA

Prof(a). Mara Abel
UFRGS

Prof(a). Tiago Prince Sales
University of Twente

Prof(a). Breno Giovanni Adaid Castro
IDP

Prof(a). Ricardo Rodrigues Barbosa
Aposentado/UFMG

Belo Horizonte, 18 de março de 2025.



Documento assinado eletronicamente por **Marcello Peixoto Bax, Professor do Magistério Superior**, em 08/05/2025, às 11:40, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Jose Eduardo Santarem Segundo, Usuário Externo**, em 12/05/2025, às 13:34, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Ana Kelly Arantes, Usuário Externo**, em 12/05/2025, às 13:40, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **José Maria Parente de Oliveira, Usuário Externo**, em 13/05/2025, às 11:00, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Breno Giovanni Adaid Castro, Usuário Externo**, em 13/05/2025, às 14:09, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Mara Abel, Usuária Externa**, em 13/05/2025, às 17:43, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Tiago Prince Sales, Usuário Externo**, em 14/05/2025, às 03:57, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Ricardo Rodrigues Barbosa, Professor Magistério Superior - Voluntário**, em 20/05/2025, às 17:26, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **4134241** e o código CRC **1AB1FFA7**.



UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE CIÊNCIA DA INFORMAÇÃO - ECI
PROGRAMA DE PÓS-GRADUAÇÃO EM GESTÃO E ORGANIZAÇÃO DO CONHECIMENTO - PPGOC

ATA DA DEFESA DE TESE DO ALUNO

EDUARDO CARDOSO MELO

Realizou-se, no dia 18 de março de 2025, às 14:00 horas, por videoconferência, da Universidade Federal de Minas Gerais, a defesa de tese, intitulada *ORIENTAÇÃO A DADOS E ONTOLOGIAS: ANÁLISE DAS RELAÇÕES ENTRE A CONCESSÃO DE AUXÍLIOS SOCIOECONÔMICOS, EVASÃO ESCOLAR E DESEMPENHO ACADÊMICO NO IFMG*, apresentada por EDUARDO CARDOSO MELO, número de registro 2021651325, graduado no curso de TECNOLOGIA EM INFORMÁTICA, como requisito parcial para a obtenção do grau de Doutor em GESTÃO E ORGANIZAÇÃO DO CONHECIMENTO, à seguinte Comissão Examinadora: Prof(a). Marcello Peixoto Bax - ECI/UFMG (Orientador), Dr(a). Ana Kelly Arantes - IFMG, Prof(a). José Eduardo Santarem Segundo - USP, Prof(a). José Maria Parente de Oliveira - ITA, Prof(a). Mara Abel - UFRGS, Prof(a). Tiago Prince Sales - University of Twente, Prof(a). Breno Giovanni Adaid Castro - IDP, Prof(a). Ricardo Rodrigues Barbosa - Aposentado/UFMG.

A Comissão considerou a tese:

Aprovada

Reprovada

Finalizados os trabalhos, lavrei a presente ata que, lida e aprovada, vai assinada por mim e pelos membros da Comissão.

Belo Horizonte, 18 de março de 2025.

Assinatura dos membros da banca examinadora:



Documento assinado eletronicamente por **Marcello Peixoto Bax, Professor do Magistério Superior**, em 08/05/2025, às 11:40, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Jose Eduardo Santarem Segundo, Usuário Externo**, em 12/05/2025, às 13:35, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Ana Kelly Arantes, Usuário Externo**, em 12/05/2025, às 13:40, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **José Maria Parente de Oliveira, Usuário Externo**, em 13/05/2025, às 10:58, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Breno Giovanni Adaid Castro, Usuário Externo**, em 13/05/2025, às 14:09, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Mara Abel, Usuária Externa**, em 13/05/2025, às 17:42, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Tiago Prince Sales, Usuário Externo**, em 14/05/2025, às 03:57, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Ricardo Rodrigues Barbosa, Professor Magistério Superior - Voluntário**, em 20/05/2025, às 17:26, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **4134230** e o código CRC **E29AD6A6**.

AGRADECIMENTOS

Agradeço, primeiramente, a Deus, por me proporcionar todas as condições para concluir este trabalho.

À minha amada esposa Fernanda, minha maior incentivadora e companheira!

Aos meus filhos, Lucca e Clara, por trazerem tanta alegria e amor para minha vida!

Aos meus pais, Elvira e Dilermando, por todo incentivo e auxílio!

Ao meu orientador, Professor Marcello Peixoto Bax, por tanta sabedoria compartilhada, paciência nas orientações e apoio ao longo de toda a jornada!

Aos colegas de curso e professores do PPGGOC, pela oportunidade de obter novos conhecimentos!

Aos colegas do IFMG e à instituição, por contribuírem diretamente para a realização deste estudo!

RESUMO

A análise do ensino superior brasileiro espelha um panorama com diversas mudanças ao longo do tempo, notadamente desde o ano 2000, onde o aumento da oferta e ampliação do acesso a diferentes cursos proporcionaram à população oportunidades de qualificação até então restritas aos indivíduos com capacidade financeira para arcar com este investimento. Dentre os desafios surgidos neste novo contexto, destaca-se a complexidade de manutenção do vínculo institucional com os estudantes, refletido nas estatísticas de evasão escolar. Considerando que a oferta de auxílios socioeconômicos foi uma das estratégias governamentais para apoiar a permanência dos estudantes, esta tese investigou, com base nos recursos conceituais e práticos de diversas áreas de conhecimento, como a Ciência da Informação, a relação da oferta de tais benefícios com a evasão escolar e o desempenho acadêmico no âmbito dos cursos de graduação ofertados por uma Instituição Federal de Ensino Superior (IFES). Trata-se de um estudo classificado como pesquisa aplicada, exploratória e explicativa, com abordagem qualitativa e quantitativa, utilizando como procedimentos a pesquisa *Ex-Post Facto*, a pesquisa experimental e o estudo de caso. Quanto à concepção filosófica, tem-se uma pesquisa categorizada como pós-positivista e pragmática. Os procedimentos metodológicos foram divididos em três etapas, dada a abrangência das atividades realizadas. O desenvolvimento da primeira etapa gerou como principais resultados um amplo levantamento bibliográfico sobre os temas pertinentes à tese, um estudo caracterizador do perfil dos candidatos a auxílios socioeconômicos na instituição pesquisada e outro estudo que construiu o perfil de comportamento informacional dos servidores que atuam nos processos seletivos para concessão dos benefícios. Na segunda etapa foi desenvolvida uma estrutura tecnológica, baseada em ontologia, para integrar os dados dos três construtos analisados na tese. A estrutura criada permitiu que, na terceira etapa, fossem elaboradas análises sobre as relações entre a concessão de auxílio socioeconômico, evasão escolar e desempenho acadêmico sob diferentes óticas e com a aplicação de técnicas complementares. Como principais resultados, destacam-se: I) a identificação de relação positiva e significativa entre os construtos, indicando que, na instituição pesquisada, a concessão do auxílio contribui para a redução das taxas de evasão e para o incremento no desempenho acadêmico dos estudantes; II) a elaboração e validação de uma ontologia para este domínio, a qual pode ser reutilizada em contextos similares por outras instituições educacionais; III) uma análise baseada em dados do impacto do Programa Nacional de Assistência Estudantil (PNAES) na instituição pesquisada. Finalmente, a tese reforça o caráter transdisciplinar da Ciência da Informação ao promover diversas interações com outras áreas de conhecimento para a geração de produtos informacionais que, por sua vez, habilitam os indivíduos a construir novos conhecimentos lastreados em dados reais e não somente em suposições subjetivas.

Palavras-chave: Assistência Estudantil. Evasão Escolar. Desempenho Acadêmico. Ontologia. Integração de Dados. Ciência de Dados.

ABSTRACT

The analysis of Brazilian higher education reflects a panorama that has changed several times over time, especially since the year 2000, when the increase in the supply and expansion of access to different courses provided the population with opportunities for qualification that were previously restricted to individuals with the financial capacity to afford this investment. Among the challenges that have arisen in this new context, the complexity of maintaining institutional ties with students stands out, reflected in school dropout statistics. Considering that the provision of socioeconomic aid was one of the government strategies to support student retention, this thesis investigated, based on conceptual and practical resources from several areas of knowledge, such as Information Science, the relationship between the provision of such benefits and school dropout and academic performance in undergraduate courses offered by a Federal Higher Education Institution (IFES). This study is classified as applied, exploratory and explanatory research, with a qualitative and quantitative approach, using Ex-Post Facto research, experimental research and case study as procedures. Regarding the philosophical conception, the research is categorized as post-positivist and pragmatic. The methodological procedures were divided into three stages, given the scope of the activities carried out. The development of the first stage generated as main results a broad bibliographic survey on the themes pertinent to the thesis, a study characterizing the profile of candidates for socioeconomic aid in the institution studied and another study that constructed the informational behavior profile of the employees who work in the selection processes for granting the benefits. In the second stage, a technological structure, based on ontology, was developed to integrate the data from the three constructs analyzed in the thesis. The structure created allowed, in the third stage, analyses to be prepared on the relationships between the granting of socioeconomic aid, school dropout and academic performance from different perspectives and with the application of complementary techniques. The main results include: I) the identification of a positive and significant relationship between the constructs, indicating that, in the institution studied, the granting of aid contributes to the reduction of dropout rates and to the increase in the academic performance of students; II) the development and validation of an ontology for this domain, which can be reused in similar contexts by other educational institutions; III) a data-based analysis of the impact of the National Student Assistance Program (PNAES) on the institution under study. Finally, the thesis reinforces the transdisciplinary nature of Information Science by promoting diverse interactions with other areas of knowledge to generate informational products that, in turn, enable individuals to construct new knowledge based on real data and not only on subjective assumptions.

Keywords: Student Assistance. School Dropout. Academic Performance. Ontology. Data Integration. Data Science.

LISTA DE FIGURAS

Figura 1 – Elementos que compõem o estudo	18
Figura 2 – Fatores que podem influenciar a evasão	30
Figura 3 – Evolução da evasão, retenção e conclusão entre 2017 e 2021	37
Figura 4 – Linha do tempo da assistência estudantil no Brasil	41
Figura 5 – Despesa corrente da assistência ao estudante de ensino superior	46
Figura 6 – Caracterização dos perfis de vulnerabilidade da PAE/IFMG	50
Figura 7 – Ciclo de gerenciamento da informação de Choo	61
Figura 8 – Esquema da metáfora do <i>sense-making</i>	68
Figura 9 – Modelo de comportamento informacional de Choo	70
Figura 10 – Processos e fases da Metodologia SABiO	81
Figura 11 – Tipos de Aprendizado de Máquina	89
Figura 12 – Exemplo de aplicação do Método Elbow	94
Figura 13 - Fases do ciclo de construção de políticas públicas	96
Figura 14 – Etapas da pesquisa	105
Figura 15 – Atividades para caracterização do perfil dos candidatos a auxílios	106
Figura 16 – Atividades para construção do ambiente integrado à ontologia	112
Figura 17 – Fontes para coleta dos dados e ponto de integração	114
Figura 18 – Solicitações de auxílios por ano do edital	122
Figura 19 – Distribuição geral das bolsas	123
Figura 20 – Solicitações por renda <i>per capita</i>	125
Figura 21 – Solicitações por modalidade do curso	127
Figura 22 – Solicitações por forma de ingresso	128
Figura 23 – Solicitações por situação de trabalho	128
Figura 24 – Solicitações de acordo com o recebimento de auxílios no ano anterior	129
Figura 25 – Idade média na inscrição por Campus	130
Figura 26 – Renda <i>per capita</i> de acordo com a situação final das solicitações	131
Figura 27 – Situação final das solicitações conforme a idade na inscrição	132
Figura 28 – Execução do Método <i>Elbow</i> no conjunto de dados	132

Figura 29 – Agrupamentos criados pelo <i>K-means</i>	133
Figura 30 – Resultados da métrica <i>Silhouette Score</i>	134
Figura 31 – Idade na inscrição nos três agrupamentos	136
Figura 32 – Renda <i>per capita</i> nos três agrupamentos	136
Figura 33 – Participantes conforme o tempo de trabalho na assistência estudantil ..	138
Figura 34 – Necessidade de informação	139
Figura 35 – Canais de informação	140
Figura 36 – Fontes de informação	141
Figura 37 – Barreiras para a busca de informação	142
Figura 38 – Sentimentos percebidos na busca de informação	142
Figura 39 – Meios de compartilhamento de informação	144
Figura 40 – Elementos considerados ao compartilhar informações	144
Figura 41 – Ontologia de referência elaborada com OntoUML	153
Figura 42 – Verificação sintática da ontologia conforme as regras da OntoUML	155
Figura 43 – Fragmento de código do arquivo com a ontologia exportada em gUFO	155
Figura 44 – Integração das classes da Onto-Educacional com a gUFO	156
Figura 45 – Propriedades dos objetos da ontologia	157
Figura 46 – Propriedades dos dados da ontologia	157
Figura 47 – Representação gráfica da ontologia operacional	158
Figura 48 – Verificação sintática da ontologia operacional conforme a UFO	159
Figura 49 – Grafo com as relações de um curso específico	169
Figura 50 – Grafo com os cursos ofertados por determinado Campus	169
Figura 51 – Grafo com os cursos de grau Tecnológico ofertados pelo IFMG	170
Figura 52 – Grafo com as relações de um estudante bolsista	171
Figura 53 – Propriedades de um nó selecionado	171
Figura 54 – Primeira aba do <i>dashboard</i>	187
Figura 55 – Segunda aba do <i>dashboard</i>	188
Figura 56 – Quarta aba do <i>dashboard</i>	190
Figura 57 – Quinta aba do <i>dashboard</i>	197
Figura 58 – Sexta aba do <i>dashboard</i>	198

Figura 59 – Sétima aba do <i>dashboard</i>	199
Figura 60 – Oitava aba do <i>dashboard</i>	200
Figura 61 – Nona aba do <i>dashboard</i>	201
Figura 62 – Histograma com os escores de propensão calculados	210
Figura 63 – Distribuição dos dados nos conjuntos completo e pareado	212
Figura 64 – Diferença do escore de propensão das variáveis independentes	212
Figura 65 – Histogramas do escore de propensão calculado	213

LISTA DE TABELAS

Tabela 1 – Situação dos estudantes de cursos superiores do IFMG (2008 a 2015)	35
Tabela 2 – Distribuição de auxílio estudantil nas IFES	46
Tabela 3 – Distribuição de bolsa permanência e recursos empenhados (em R\$) ...	51
Tabela 4 – Variáveis relacionadas com o desempenho acadêmico	55
Tabela 5 – Atributos dos candidatos a auxílios socioeconômicos	120
Tabela 6 – Solicitações de auxílios por Campus	124
Tabela 7 – Cinco cursos com mais solicitações de bolsas	125
Tabela 8 – Cinco cursos com menos solicitações de bolsas	126
Tabela 9 – Importância dos atributos	131
Tabela 10 – Tipos de informação	140
Tabela 11 – Usos da informação	143
Tabela 12 – Importância das características da informação	145
Tabela 13 – Resultados da consulta para a questão de competência 1	173
Tabela 14 – Validação estatística dos resultados da questão de competência 1	173
Tabela 15 – Resultados da consulta para a questão de competência 2	175
Tabela 16 – Resultados da consulta para a questão de competência 3	176
Tabela 17 – Resultados da consulta para a questão de competência 4 (gênero)	177
Tabela 18 – Validação estatística da questão de competência 4 (gênero)	178
Tabela 19 – Resultados da consulta para a questão de competência 4 (cor)	178
Tabela 20 – Validação estatística da questão de competência 4 (cor)	179
Tabela 21 – Resultados da consulta para a questão de competência 5	181
Tabela 22 – Resultados da consulta para a questão de competência 6	182
Tabela 23 – Resultados da consulta para a questão de competência 7	183
Tabela 24 – Resultados da consulta para a questão de competência 8	185
Tabela 25 – Relação entre os três construtos por campus	189
Tabela 26 – Evasão por curso e turmas nos <i>campi</i> que ofertam apenas um curso ..	191
Tabela 27 – Evasão por curso e turmas no Campus Bambuí	191
Tabela 28 – Evasão por curso e turmas no Campus Betim	192

Tabela 29 – Evasão por curso e turmas no Campus Congonhas	192
Tabela 30 – Evasão por curso e turmas no Campus Formiga	193
Tabela 31 – Evasão por curso e turmas no Campus Governador Valadares	193
Tabela 32 – Evasão por curso e turmas no Campus Ouro Branco	194
Tabela 33 – Evasão por curso e turmas no Campus Ouro Preto	194
Tabela 34 – Evasão por curso e turmas no Campus Ribeirão das Neves	195
Tabela 35 – Evasão por curso e turmas no Campus Sabará	195
Tabela 36 – Evasão por curso e turmas no Campus Santa Luzia	196
Tabela 37 – Evasão por curso e turmas no Campus São João Evangelista	196
Tabela 38 – Proporção de bolsistas, evadidos e formados por campus	197
Tabela 39 – Médias das variáveis nos dois grupos	208
Tabela 40 – Variáveis independentes preditoras	209
Tabela 41 – Comparação das variáveis nos grupos de tratamento e controle	211
Tabela 42 – Impacto do auxílio socioeconômico sobre desempenho e evasão	214

LISTA DE QUADROS

Quadro 1 – Cinco atributos mais importantes de cada componente	133
Quadro 2 – Caracterização dos agrupamentos	135
Quadro 3 – Sumarização dos resultados sobre Comportamento Informacional	146
Quadro 4 – Definição dos conceitos da ontologia	151
Quadro 5 – Atributos mapeados conforme a ontologia operacional	163
Quadro 6 – Matriz de rastreabilidade das questões de competência da ontologia	186
Quadro 7 – Atributos do conjunto de dados para execução do PSM	206

LISTA DE ABREVIATURAS

AM	–	Aprendizado de Máquina
BI	–	<i>Business Intelligence</i>
BP	–	Bolsa Permanência
CAE	–	Coordenação de Assistência Estudantil
CI	–	Ciência da Informação
CR	–	Coeficiente de Rendimento
CSV	–	<i>Comma-Separated Values</i>
DIRAE	–	Diretoria de Assuntos Estudantis
ENEM	–	Exame Nacional do Ensino Médio
GC	–	Gestão do Conhecimento
GI	–	Gestão da Informação
IA	–	Inteligência Artificial
IES	–	Instituição de Ensino Superior
IFES	–	Instituição Federal de Ensino Superior
IFMG	–	Instituto Federal de Minas Gerais
IVS	–	Índice de Vulnerabilidade Social
MEC	–	Ministério da Educação
NASIFMG	–	Núcleo de Assistentes Sociais do IFMG
OWL	–	<i>Ontology Web Language</i>
PAE	–	Política de Assistência Estudantil
PCA	–	<i>Principal Component Analysis</i>
PNAES	–	Programa Nacional de Assistência Estudantil
RDF	–	<i>Resource Description Framework</i>
RDFS	–	<i>RDF Schema</i>
REUNI	–	Programa de Reestruturação e Expansão das Universidades Federais
RPC	–	Renda <i>per capita</i>
SABiO	–	<i>Systematic Approach for Building Ontologies</i>
SGBD	–	Sistema Gerenciador de Banco de Dados
SISU	–	Sistema de Seleção Unificada
SQL	–	<i>Structured Query Language</i>
TI	–	Tecnologia da Informação
UFO	–	<i>Unified Foundational Ontology</i>
UML	–	<i>Unified Modeling Language</i>
XML	–	<i>eXtensible Markup Language</i>

SUMÁRIO

1 INTRODUÇÃO	15
1.1 PROBLEMA DE PESQUISA	19
1.2 OBJETIVOS	21
1.3.1 Objetivo geral	21
1.3.2 Objetivos específicos.....	21
1.3 JUSTIFICATIVA E RELEVÂNCIA	22
1.4 CONTRIBUIÇÕES PRINCIPAIS	24
1.5 USO DE FERRAMENTAS DE INTELIGÊNCIA ARTIFICIAL	25
1.6 ORGANIZAÇÃO GERAL DA TESE	25
2 REFERENCIAL TEÓRICO	26
2.1 EVASÃO ESCOLAR	26
2.1.1 Causas e consequências da evasão	29
2.1.2 A evasão no IFMG	35
2.2 ASSISTÊNCIA ESTUDANTIL	38
2.2.1 Caracterização da assistência estudantil no IFMG	47
2.3 DESEMPENHO ACADÊMICO	52
2.3.1 Determinantes do desempenho acadêmico	54
2.4 GESTÃO DA INFORMAÇÃO (GI)	57
2.4.1 Relações entre a Tecnologia da Informação e a GI	62
2.4.2 Comportamento informacional	65
2.5 ONTOLOGIAS	72
2.5.1 Unified Foundational Ontology (UFO)	76
2.5.2 Construção e representação de ontologias	77
2.5.3 Integração semântica de dados baseada em ontologia	79
2.5.4 Abordagem de engenharia de ontologias (SABiO)	80
2.6 CIÊNCIA DE DADOS	83
2.6.1 Produção e armazenamento dos dados	84

2.6.2	Preparação e transformação dos dados	85
2.6.3	Análise de dados	87
2.6.3.1	Agrupamento de instâncias	91
2.7	AVALIAÇÃO DE POLÍTICAS PÚBLICAS	94
2.8	TRABALHOS CORRELATOS	98
3	METODOLOGIA	102
3.1	CARACTERIZAÇÃO DA PESQUISA	102
3.2	PROCEDIMENTOS METODOLÓGICOS	105
3.2.1	Etapa 1	105
3.2.1.1	Realização de levantamentos bibliográficos	105
3.2.1.2	Caracterização do perfil dos candidatos a auxílio socioeconômico	106
3.2.1.3	Caracterização do comportamento informacional dos servidores	108
3.2.2	Etapa 2	110
3.2.2.1	Construção e avaliação da ontologia de domínio	110
3.2.3	Etapa 3	116
3.2.3.1	Análise das relações entre os construtos e avaliação da ontologia	116
3.2.3.2	Análise sobre a efetividade do PNAES na instituição	117
4	RESULTADOS E DISCUSSÕES	119
4.1	CARACTERIZAÇÃO DO PERFIL DOS CANDIDATOS EM PROCESSOS SELETIVOS DA ASSISTÊNCIA ESTUDANTIL	119
4.2	CARACTERIZAÇÃO DO PERFIL DE COMPORTAMENTO INFORMACIONAL DOS SERVIDORES ENVOLVIDOS COM O PROCESSO DE CONCESSÃO DE AUXÍLIO SOCIOECONÔMICO	137
4.3	CONSTRUÇÃO E AVALIAÇÃO DA ONTOLOGIA DE DOMÍNIO	147
4.3.1	Identificação de propósito e definição de requisitos	148
4.3.2	Captura e formalização da ontologia	149
4.3.3	Design, implementação e testes	155
4.3.3.1	Design e implementação da ontologia operacional	155
4.3.3.2	Coleta e construção do conjunto de dados	159
4.3.3.3	Ingestão dos dados em um <i>Graph Database</i>	167

4.3.3.4 Avaliação da ontologia	172
4.4 ANÁLISE DAS RELAÇÕES ENTRE OS CONSTRUTOS	186
4.5 ANÁLISE SOBRE A EFETIVIDADE DO PNAES NA INSTITUIÇÃO	206
5 CONSIDERAÇÕES FINAIS	215
5.1 CONSIDERAÇÕES QUANTO AOS OBJETIVOS ESPECÍFICOS	216
5.2 CONSIDERAÇÕES QUANTO AOS OBJETIVOS GERAIS	218
5.3 CONSIDERAÇÕES PARA A ÁREA DE CIÊNCIA DA INFORMAÇÃO	222
5.4 DIFICULDADES ENCONTRADAS E TRABALHOS FUTUROS	224
5.5 PUBLICAÇÕES DOS RESULTADOS OBTIDOS	225
REFERÊNCIAS	228
APÊNDICE A – PESQUISA SOBRE O COMPORTAMENTO INFORMACIONAL .	264

1 INTRODUÇÃO

A educação superior no Brasil passou por expressivas transformações desde o ano 2000, tanto em termos de oferta quanto em relação ao perfil dos alunos. Se antes as universidades e faculdades ficavam localizadas em grandes centros, fazendo com que os alunos tivessem que se mudar durante os estudos, a expansão do ensino privado e do público levaram a uma interiorização das instituições e, conseqüentemente, criaram novas possibilidades para grupos de pessoas que não teriam condições de arcar com todos os custos que envolvem esta mudança ([CAMARGO e ARAÚJO, 2018](#)). No setor público, o Programa de Reestruturação e Expansão das Universidades Federais (Reuni) é considerado um dos principais responsáveis pelo crescimento do número de Instituições Federais de Ensino Superior (IFES), *campi*, cursos e vagas no Brasil ([TREVIZAN, 2019](#)). Dentre seus resultados, destaca-se que entre 2003 e 2011 houve um crescimento de aproximadamente 111% no número de vagas ofertadas em universidades federais (de 109.184 para 231.530), além de quase 100% de incremento do número de estudantes matriculados, passando de 596.219 para 1.029.141 ([BRASIL, 2012a](#)).

Outras ações governamentais, especialmente aquelas caracterizadas como políticas públicas, também contribuíram no incremento da oferta de ensino superior para a sociedade brasileira, dentre as quais se destacam a Lei das Cotas (Nº 12.711, de 29 de agosto de 2012), o Sistema de Seleção Unificada (SISU), o PROUNI (Programa Universidade Para Todos), o Programa Nacional de Assistência Estudantil (PNAES) e a criação dos Institutos Federais de Educação, Ciência e Tecnologia (Lei Nº 11.892, de 29 de dezembro de 2008). Dois indicadores de 2020 mostram, sob a ótica quantitativa, os resultados de tais ações no que se refere à ampliação do acesso ao ensino superior: o número de vagas ofertadas pelas IFES para cursos de graduação foi de 492.599 (com ocupação de 80%) e a quantidade de matrículas foi de 1.254.080 ([INEP, 2022](#)). O fomento à continuidade de tais políticas, bem como a proposição de novos programas que sirvam para complementar aspectos não contemplados ou atendidos parcialmente, precisam ser lastreados em robustas avaliações dos impactos e resultados das políticas implementadas pelo governo ([CUNHA, 2018](#)). Essas avaliações fornecem evidências empíricas sobre a efetividade das políticas e contribuem para a prestação de contas, tomada de decisão, aprendizado organizacional e eficiência na alocação de recursos ([PINHEIRO, 2020](#)).

A concepção dos Institutos Federais de Educação, Ciência e Tecnologia ocorreu neste momento de valorização do ensino público, aumentando a capilaridade e alcance da rede federal de educação profissional para os mais remotos locais do país em mais de 685 unidades vinculadas a quase 40 Institutos Federais ([MEC, 2024](#)). A criação do Instituto Federal de Minas Gerais (IFMG) aconteceu em 2008, tendo uma Reitoria sediada em Belo

Horizonte e dezoito *campi* localizados, em sua maioria, na região central do estado. São ofertados cursos técnicos (integrados, subsequentes e concomitantes), graduações (bacharelados, licenciaturas e tecnológicos) e pós-graduações (*Lato Sensu* e *Stricto Sensu*), objetivando promover a verticalização do ensino para mais de vinte mil estudantes matriculados ([BRASIL, 2024](#); [IFMG, 2022a](#)). É válido ressaltar que os institutos federais têm foco na educação técnica e tecnológica, com forte integração entre ensino médio e superior, enquanto as universidades federais são voltadas majoritariamente para a formação acadêmica e pesquisa em nível superior e de pós-graduação.

O aumento da oferta de cursos e vagas provocou outras questões relacionadas com a gestão das IFES, em especial aquelas que dizem respeito à evasão dos estudantes e o seu desempenho acadêmico. Ao possibilitar o ingresso de indivíduos com perfis e realidades distintas das vigentes até então, a expansão da rede federal fez com que novas variáveis tivessem influência na descontinuidade do curso e no desempenho dos estudantes, tais como renda mensal familiar *per capita*, condições de moradia e transporte até o campus, recursos para aquisição de alimentos e materiais de estudo, dentre outras ([NONATO et al. 2020](#); [RISTOFF, 2014](#); [SENKEVICS e MELLO, 2019](#)). A pesquisa realizada por [ANDIFES \(2019\)](#) apresenta um perfil variado do atual estudante de graduação em IFES no país, o que sugere a necessidade de adaptação da gestão pública para entender os fatores que impactam na evasão e no desempenho deste público.

Uma das ações realizadas pelas IFES com o objetivo de possibilitar a manutenção dos estudantes no ambiente acadêmico, bem como contribuir para o seu desempenho ao longo do curso, foi o incremento na oferta de programas de assistência estudantil para os mais vulneráveis tanto econômica quanto socialmente. O apoio de tais programas seria uma estratégia eficaz para a permanência discente e para a obtenção de resultados satisfatórios nas atividades previstas pelo curso ([SANTOS, 2020](#)). Visando fortalecer em toda a rede federal as políticas de assistência estudantil, bem como oferecer condições para que os estudantes permaneçam na instituição até a conclusão do seu curso, o Governo Federal criou em 2007 o PNAES. Em 2010 o Programa foi institucionalizado e publicado como [Decreto](#), indicando a necessidade de atenção para com moradia, alimentação, transporte e apoio pedagógico aos estudantes ([BRASIL, 2010](#)). Desde então, as IFES passaram a contar com uma orientação formal que fomentou a atuação da assistência estudantil mediante dotação orçamentária específica ([PINHEIRO, 2018](#)). Em 2024 o Governo Federal promoveu a transformação do PNAES em uma política nacional por meio da Lei Nº 14.914, de 3 de julho de 2024 ([BRASIL, 2024](#)).

O PNAES requer a realização de processos seletivos com critérios e metodologias para a definição de quais discentes serão contemplados. Além do atendimento a tais critérios, os candidatos precisam possuir renda familiar *per capita* de até um salário mínimo

([BRASIL, 2024](#)). Uma forma de coletar os dados dos interessados é o preenchimento de questionários socioeconômicos e a entrega de documentação comprobatória. A consolidação desses dados permite realizar uma análise mais detalhada da realidade de cada estudante e subsidiar a tomada de decisão para atribuir ou não determinado auxílio. Contudo, dependendo da amplitude da coleta de dados, da estrutura organizacional da IFES e da disponibilidade de recursos computacionais e de pessoal, a análise pode ficar prejudicada e impactar na não atribuição de auxílios para aqueles mais necessitados ([REIS e GONÇALVES, 2018](#)).

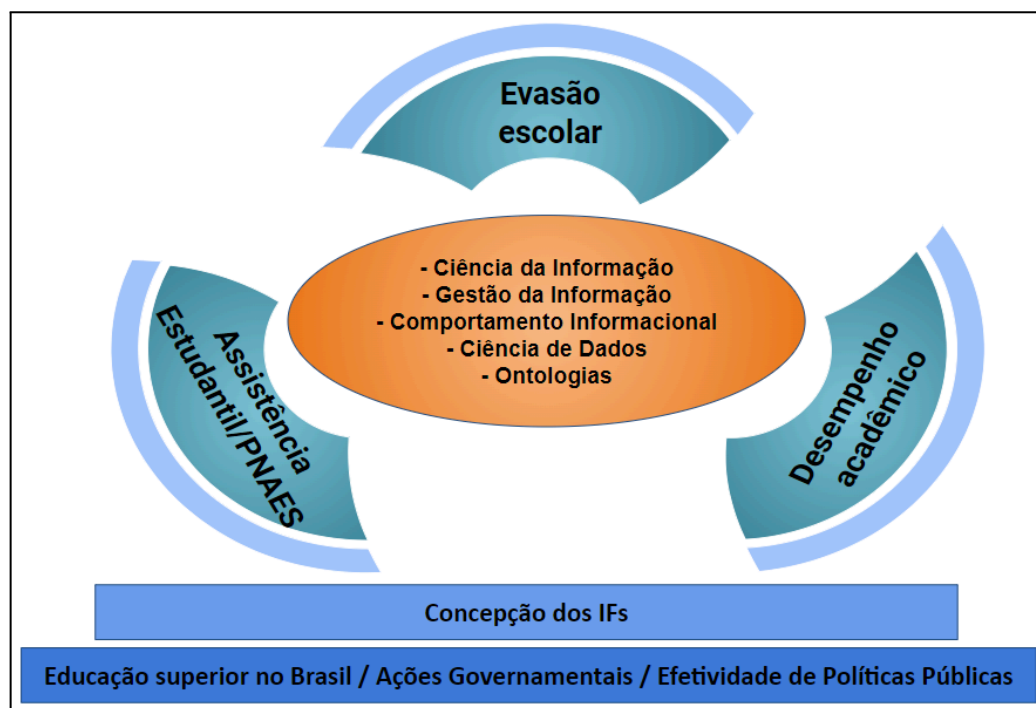
Esta preocupação com eventuais dificuldades no tratamento dos dados coletados é uma das consequências do fato de a sociedade moderna produzir uma quantidade enorme de dados e necessitar que eles sejam transformados em informações úteis para processos decisórios. Quando esta situação é analisada do ponto de vista organizacional, onde as decisões são capazes de afetar direta ou indiretamente um grande número de pessoas, dispor de uma gestão consistente dos dados e informações é fundamental para que os resultados sejam satisfatórios ([SILVA e TOMÁEL, 2007](#)).

Neste contexto, a Gestão da Informação (GI) é uma área que atua no gerenciamento dos diversos sistemas e processos existentes nas organizações utilizados em tarefas que envolvem dados e informações, tais como criação, aquisição, organização, armazenamento, distribuição e utilização. Seu objetivo é auxiliar as pessoas a acessarem informações úteis para diversos fins, especialmente aqueles relacionados com etapas de processos decisórios ([DETLOR, 2010](#)). Outra atribuição da GI é fornecer meios para que seja possível a integração dos dados de diferentes sistemas. [Barbosa e Bax \(2016\)](#) reforçam este entendimento ao tratarem as novas demandas de sistemas que tenham habilidades de integração como exigência, sob pena de ser criado um ambiente no qual cada sistema resolve problemas específicos e opera de forma isolada. Em um mundo amplamente dependente de recursos tecnológicos, a todo momento coletando e intercambiando dados, a Gestão da Informação mostra-se essencial.

As ontologias, conceituadas por [Borst \(1997, p. 12\)](#) como a “especificação formal de uma conceitualização compartilhada”, são entendidas como recursos eficazes para auxiliar neste processo de integração e interoperabilidade semântica de dados oriundos de diferentes fontes ([CAMPOS et al., 2020](#); [CARCHEDI et al., 2018](#); [GONÇALVES, 2020](#)). Com a proposta de elaboração de um modelo comum para os dados, as ontologias conseguem produzir uma representação mais consistente sobre o conhecimento de determinado domínio a partir da formalização dos conceitos e relações existentes no mesmo, de maneira que os dados envolvidos estejam em um formato anotado semanticamente e legível por máquinas ([DE GIACOMO et al., 2018](#)).

Essa contextualização pode ser resumida graficamente na Figura 1, onde tem-se a integração de todos os elementos que compreendem o escopo desta tese.

Figura 1: Elementos que compõem o estudo



Fonte: Elaborada pelo autor (2023)

A base do estudo reside no contexto da educação superior brasileira, a qual é suportada por ações governamentais e políticas públicas formalizadas que requerem constantes análises de efetividade/impacto quanto ao alcance dos seus objetivos. Os Institutos Federais de Educação, Ciência e Tecnologia foram criados em um momento de expansão do ensino público federal e também fazem parte do escopo da tese enquanto objeto do estudo de caso. Os três construtos cujas análises são suportadas por uma base ontológica envolvem a **evasão escolar**, o **desempenho acadêmico** e a **concessão de auxílio socioeconômico** por programas de assistência estudantil (em consonância com o preconizado pelo PNAES). A formulação de conclusões acerca das relações entre os três construtos no âmbito do IFMG, envolvendo para isso diversas áreas de conhecimento (Ciência da Informação, Gestão da Informação, Comportamento Informacional, Ciência de Dados e Ontologias), trouxe novas perspectivas sobre esse contexto, as quais são compartilhadas como principais resultados da tese.

1.1 Problema de pesquisa

O ensino superior nas IFES chegou a um cenário onde existe considerável disponibilidade de vagas e alto número de matrículas efetivadas, porém com índice de evasão preocupante. De acordo com [INEP \(2022\)](#), a taxa de desistência acumulada das IFES alcançou a média de 55% em 2020, contra 43% de concluintes. No âmbito do IFMG, a Plataforma Nilo Peçanha¹ reporta que, ao final de 2023, a evasão acumulada média nos cursos de graduação presencial foi de 53% no ciclo dos cursos analisados, com apenas 19% de concluintes. Isso indica, ainda, que 28% dos estudantes analisados terminaram 2023 vinculados com a instituição, mas não concluíram seu curso no tempo previsto para integralização, possivelmente em função de desempenho acadêmico insuficiente para cumprimento dos componentes curriculares. Não existia, até a elaboração desta tese, nenhum estudo publicado analisando a evasão e o desempenho acadêmico em todos os cursos de graduação ofertados pelo IFMG, somente pesquisas restritas a cursos específicos.

Partindo do princípio que o PNAES deveria contribuir para a redução da evasão escolar e melhora no desempenho acadêmico nas IFES, a existência de um monitoramento contínuo mensurando e analisando o seu impacto nesses aspectos seria benéfica para todos os envolvidos (IFES, Governo Federal, sociedade civil, dentre outros). Entretanto, esta ainda não é a realidade encontrada atualmente ([BRITO, COSTA e ALMEIDA, 2019](#); [FAVA, HALL e CINTRA, 2022](#); [SANTOS, SIQUEIRA e DINIZ, 2021](#)). Por parte do Governo Federal, o PNAES não tem seus resultados medidos e divulgados com a devida periodicidade ([GUIMARÃES, 2020](#)). Estudos que visam promover uma análise do PNAES muitas vezes o fazem sob o ponto de vista descritivo ([BRAGA e PRÁ, 2021](#); [ELOI et al., 2019](#); [FREITAS, 2021](#); [GOMES e PASSOS, 2018](#); [RECKTENVALD, MATTEI e PEREIRA, 2018](#)), buscando explorar como as ações da assistência estudantil são conduzidas pela IFES, sem focar no alcance dos resultados e no impacto promovido pelas ações do programa.

Como cada IFES está habilitada pelo PNAES a definir seu próprio conjunto de regras e critérios para seleção dos estudantes aptos ao recebimento dos auxílios socioeconômicos, cenários variados podem ser construídos onde o volume de dados coletados dos candidatos é desproporcional ao necessário para o processo de análise e deliberação, isto é, são coletadas e armazenadas informações que, eventualmente, não teriam utilidade sob o ponto de vista do processo decisório ([BROWN, 2020](#); [DIAZ e WATTS, 2019](#)). A opção por uma modelagem ontológica mais condizente com a realidade, nos termos de um conjunto mínimo de atributos que caracterizam satisfatoriamente os conceitos e relacionamentos do domínio, pode ser fundamental para uma adequada estrutura de dados que visa servir de base em decisões tão importantes como a atribuição de auxílios estudantis ([RIBEIRO, 2010](#)).

¹ <http://plataformanilopecanha.mec.gov.br/>

Mesmo em um panorama onde o volume de dados é aceitável sob o ponto de vista das estruturas de armazenamento e processamento, a quantidade de variáveis a serem analisadas faz com que o processo não seja trivial para o ser humano, demandando suporte baseado em recursos da Tecnologia da Informação (TI) ([PIETSCH, 2021](#)). O Instituto Federal de Minas Gerais define os estudantes contemplados a partir da análise de dez indicadores ([IFMG, 2020](#)). Dadas as restrições da mente humana em considerar a amplitude de ligações entre esses indicadores para cada candidato ([COSTA, 2012](#); [RABELO e GOES, 2019](#)), recursos computacionais são úteis para apoiar as rotinas envolvidas com o entendimento do perfil dos candidatos e com a análise socioeconômica em geral.

Outra dificuldade enfrentada pelas instituições de ensino refere-se à integração de dados oriundos de fontes heterogêneas e com formatos diferentes ([CARCHEDI et al., 2018](#); [GARCIA, ABÍLIO e MALHEIROS, 2015](#); [JORGE et al., 2020](#); [KHAMDAMOV, 2021](#)), fazendo com que a geração de novos conhecimentos fique prejudicada em função de um ambiente incapaz de fornecer os subsídios adequados para o entendimento de questões acadêmicas ([RODZI, OTHMAN e YUSUF, 2015](#)). Os motivos são diversos, tais como a utilização de softwares fornecidos por empresas diferentes ([GHELLER et al., 2017](#)), disponibilidade de dados em bases governamentais sem padronização ([CABRAL, 2021](#); [FONSECA, AZEVEDO e ALMEIDA, 2014](#)) e tecnologias escolhidas para estruturação e armazenamento dos dados ([ARGENTON, 2020](#)).

Trata-se de problema não exclusivo da área educacional pública, mas em um contexto marcado pela alta evasão escolar, o contorno da situação se delinea negativamente. A partir do momento que a integração dos dados é frágil ou o custo para alcançá-la é inviável, o gestor público pode ficar refém de informações que não contribuem satisfatoriamente para a tomada de decisões com qualidade ([SANNA, 2014](#)). A construção de ambientes baseados em múltiplos provedores de serviços na nuvem (*multicloud*) impõe desafios adicionais à integração de dados, devido à heterogeneidade das interfaces usadas para a troca de dados, bem como os formatos e políticas de cada provedor ([PUTRAMA e MARTINEK, 2024](#)).

A compreensão de aspectos sobre como os profissionais atuantes em setores de assistência estudantil demandam, buscam e consomem informações possui importância para que as instituições tenham a possibilidade de desenvolver os mecanismos necessários ao seu atendimento, porém poucos são os estudos que objetivam levantar e analisar o perfil de comportamento informacional destes profissionais, como encontrado em [Souza \(2020\)](#). A partir do momento em que essa caracterização não é promovida e divulgada adequadamente, corre-se o risco de ocasionar problemas no processo de concessão dos auxílios socioeconômicos, que é suportado, em grande parte, pela análise de informações dos candidatos pelos profissionais da assistência estudantil.

Considerando esta problematização e as lacunas que ainda poderiam ser exploradas, além da intenção de envolver o estudo das relações entre os três construtos, a presente investigação foi norteada pela seguinte questão de pesquisa: A abordagem ontológica auxilia na organização de uma estrutura informacional que permite identificar eventuais relações entre a distribuição de auxílio socioeconômico, evasão escolar e desempenho acadêmico?

1.2 Objetivos

1.2.1 Objetivo geral

Em função da abrangência da questão de pesquisa que direcionou o desenvolvimento da tese, optou-se pela definição de dois objetivos gerais separados por enfoque, mas invariavelmente conectados:

- Objetivo geral “socioeconômico”: Analisar aspectos do processo de distribuição de auxílio socioeconômico pela assistência estudantil e seu impacto no desempenho acadêmico e na evasão escolar de estudantes do IFMG, gerando conhecimentos para refletir sobre a efetividade da política pública PNAES no âmbito desta instituição.
- Objetivo geral “técnico”: Analisar como a abordagem ontológica contribui para a qualidade do processo de integração de dados advindos de fontes heterogêneas.

1.2.2 Objetivos específicos

Para compreensão adequada do escopo dos objetivos específicos definidos, é importante salientar que eles se referem somente ao contexto do IFMG em função do estudo de caso conduzido na tese.

- Caracterizar o perfil dos candidatos de processos seletivos da assistência estudantil e construir um modelo de clusterização para identificação de grupos entre eles;
- Caracterizar o perfil de comportamento informacional dos servidores envolvidos com o processo de concessão de auxílio socioeconômico;
- Construir e avaliar uma ontologia de domínio capaz de organizar e integrar os dados sobre evasão escolar, desempenho acadêmico dos estudantes e concessão de auxílio socioeconômico;
- Construir um conjunto de dados consistentes representando as instâncias da ontologia;

- Prototipar um *dashboard* gerencial com recursos de visualização de informações envolvendo os três construtos envolvidos;
- Analisar a existência de relações entre os construtos, pontuando aspectos do impacto do PNAES enquanto política pública.

1.3 Justificativa e relevância

A aplicabilidade de tal estudo se justifica sob variados aspectos. O primeiro deles está vinculado com a demanda de análises que propõem uma abordagem diferenciada para integrar e interoperar os dados dos construtos envolvidos, ancorada na construção de uma ontologia que una semanticamente e formalmente os dados destes domínios, bem como padronize o entendimento dos conceitos e relações existentes neste contexto. [Cheng, Du e Ma \(2008\)](#) indicam que a opção por uma ontologia como meio de representação formal da realidade de um domínio, pode servir de ponto de apoio ao desenvolvimento de novas aplicações semanticamente enriquecidas no ambiente das organizações. [Campos et al. \(2020\)](#) ressaltam que, sob o ponto de vista das publicações científicas mais recentes, a abordagem baseada em ontologias vem obtendo destaque quando os problemas estudados demandam integração de dados (por conseguinte, remoção de ambiguidades e subjetividades no momento da padronização de conceitos e relacionamentos advindos dos domínios compreendidos), assim evidenciando sua atual relevância para o meio acadêmico e profissional.

A integração de dados de diferentes fontes é um trabalho muitas vezes complexo ([CARCHEDI et al., 2018](#)), posto que envolve não apenas a questão dos sistemas de informação utilizados pelas instituições, mas também o formato de armazenamento definido para cada conjunto de dados. Quando as fontes internas da instituição precisam ser integradas com fontes oriundas de ambientes abertos, como no caso dos dados publicados por órgãos da Administração Pública, esta tarefa se torna ainda mais complexa, pois não necessariamente existe padronização acertada entre esses órgãos nem a utilização de formatos abertos ([CABRAL, 2021](#); [NASCIMENTO, 2018](#); [TOSIN, 2016](#)). Ao fazer uso de uma ontologia como base para a integração dos dados de evasão escolar, desempenho acadêmico e concessão de auxílio socioeconômico, a pesquisa construída nesta tese mostra-se relevante por utilizar uma abordagem que tem obtido resultados satisfatórios em termos tecnológicos e de negócio, bem como pela documentação e construção de uma série de recursos que podem ser reaproveitados por instituições de ensino superior.

A subseção 5.2 da tese dispõe de uma análise elaborada pelo pesquisador acerca da sua impressão quanto ao uso da abordagem ontológica para resolver o problema da integração dos dados dos três construtos investigados na pesquisa, indicando sua relevância para cinco pontos em especial: formalização compartilhada dos conceitos

envolvidos com o domínio trabalhado; facilitação na interpretação dos elementos que compõem a ontologia por meio da sua representação gráfica; flexibilidade para representar e implementar a integração de dados próprios da instituição com dados advindos de fontes externas; recuperação simplificada de informações com a escrita de consultas intuitivas; flexibilidade do modelo ontológico para a construção de um único conjunto de dados habilitado a múltiplas utilizações.

Em um mundo onde a geração de informações acontece em larga escala e, muitas vezes, de maneira dispersa dentro das organizações, o entendimento sobre o universo informacional dos profissionais da área de assistência estudantil permite que estratégias sejam concebidas no sentido de aprimorar o processo de concessão dos auxílios socioeconômicos em conformidade com o perfil dos candidatos. Uma análise bem fundamentada dos dados de cada candidato pode fazer com que a assistência financeira seja direcionada para alunos que precisam do auxílio para a continuidade dos seus estudos.

Outro fator a ser considerado é o orçamento destinado pelo Governo Federal para a assistência estudantil nas IFES, tendo alcançado mais de um bilhão de reais efetivamente pagos em 2020, conforme [Mussliner et al. \(2021\)](#). Mesmo sendo um alto valor absoluto, quando dividido pelas diversas ações assistenciais das IFES e entre o grande número de estudantes interessados, nem todos que necessitam do auxílio são contemplados e, eventualmente, podem não permanecer na instituição. Ademais, o PNAES indica claramente que, dentre seus objetivos, está a redução da evasão escolar (com o conseqüente aumento da permanência) e a melhora no desempenho acadêmico dos estudantes assistidos. Portanto, estudos como o realizado nessa tese fornecem subsídios para compreender se os gastos do Governo Federal com assistência estudantil produzem os efeitos esperados para o público-alvo e se estão em consonância com o que fora planejado anteriormente ([OLIVEIRA, 2019](#)).

O cenário do ensino superior em algumas IFES após o início da pandemia de COVID-19, quando houve aumento na evasão escolar, como relatado por [Amaral et al. \(2022\)](#), [Ferreira \(2021\)](#) e [Gutierre \(2021\)](#), trouxe à tona a necessidade premente da elaboração de estudos que abordem aspectos sobre este fenômeno que tantos prejuízos ocasiona para os indivíduos ([LOURENÇO et al., 2013](#)) e ao sistema educacional como um todo ([SILVA FILHO et al., 2007](#)). Quando um estudante encerra precocemente seu vínculo com uma instituição pública, a perda é de toda a sociedade, haja vista que são os recursos públicos advindos da cobrança de impostos que custeiam a oferta dos cursos, bem como a manutenção de toda a estrutura acadêmica e administrativa necessárias para seu funcionamento. Por ser multicausal, a evasão precisa ser continuamente estudada sob os mais diversos enfoques, de maneira a contribuir para que ações possam ser planejadas pelas IFES com base em informações íntegras e de qualidade.

1.4 Contribuições principais

A contribuição mais significativa desta tese consiste na integração com suporte ontológico dos dados de concessão de auxílio socioeconômico, evasão escolar e desempenho acadêmico dos estudantes de uma IFES. O ineditismo da aplicação da abordagem ontológica em conjunto com técnicas de Ciência de Dados no contexto desses construtos representa o grande diferencial inovador da pesquisa, promovendo contribuições diretas para o aspecto da transdisciplinaridade existente na área de Ciência da Informação.

Em relação ao estado da arte, a tese avança no sentido de apresentar, com detalhes, um *pipeline* para a execução de todo o ciclo de vida de construção de uma ontologia de referência, sua transformação em ontologia operacional e o uso integrado dos dados dos três construtos em uma plataforma computacional com recursos de visualização de informações. O estado da arte da área de políticas públicas também é engrandecido, pois estudos sobre o impacto do PNAES são, em sua maioria, descritivos das ações assistenciais das IFES, ao contrário do apresentado na tese, orientado a dados que permitem analisar, medir e validar estatisticamente o impacto da concessão de auxílio socioeconômico sobre a evasão escolar e o desempenho acadêmico dos estudantes.

Considerando o escopo das atividades da tese, ela contribui diretamente com a instituição pesquisada ao promover análises com dados reais obtidos em seu ambiente, todas inéditas até então: construção do perfil dos estudantes que pleitearam auxílio socioeconômico e do perfil de comportamento informacional dos servidores envolvidos com processos seletivos da assistência estudantil; resultados da evasão de todos os cursos ofertados nos *campi* e do desempenho acadêmico dos estudantes; comparações variadas envolvendo bolsistas e não bolsistas. Com a disponibilização de um ambiente tecnológico capaz de contribuir para o processo decisório acadêmico, a gestão do IFMG tende a se beneficiar dos resultados desta tese não somente na investigação dos dados históricos coletados, mas sobretudo com vistas ao planejamento de programas e ações que permitam melhorar seus indicadores vinculados aos três construtos pesquisados.

Reforçando os benefícios desta tese para a área de Ciência da Informação (CI), tem-se a entrega de produtos documentais e tecnológicos desenvolvidos à luz dos recursos conceituais disponíveis por este campo de conhecimento, bem como a difusão da integração semântica de dados com suporte ontológico, ampliando assim o caráter transdisciplinar da Ciência da Informação ([DAHER JÚNIOR e BORGES, 2021](#); [SOUZA, 2012](#)). Portanto, a pesquisa está em sintonia com o entendimento de [Barreto \(2012\)](#), quando ele trata que os produtos ou serviços de informação somente podem ser objetos de reflexão e, eventualmente, contribuir para a área de CI, quando está claro o seu vínculo com a geração de conhecimento.

As Seções 5.2 e 5.3 exploram as contribuições da tese sob aspectos adicionais e com maior profundidade de análise.

1.5 Uso de ferramentas de Inteligência Artificial

A redação deste texto foi realizada pelo autor, sem contribuição de ferramentas de Inteligência Artificial na produção do conteúdo.

1.6 Organização geral da tese

Esta tese está estruturada da seguinte forma: no Capítulo 1 foi promovida uma contextualização do tema e de todos os assuntos que o integram, abordando os problemas levantados, os objetivos propostos e a justificativa para continuidade do estudo; no Capítulo 2 é apresentado o arcabouço teórico e os trabalhos relacionados com o tema em questão; no Capítulo 3 é feita a classificação metodológica e são explicados os procedimentos adotados para a realização das atividades que, por sua vez, cumpriram os objetivos marcados; no Capítulo 4 são descritos as investigações conduzidas ao longo da pesquisa, juntamente com análise dos resultados obtidos e a discussão acerca de sua adequação frente ao que era esperado; no Capítulo 5 são feitas as considerações da tese, propostas novas possibilidades de investigações futuras e delineadas as limitações encontradas no decorrer das atividades.

2 REFERENCIAL TEÓRICO

2.1 Evasão escolar

A evasão escolar, no âmbito do ensino superior, vem sendo objeto de análises que buscam compreender os fatores geralmente associados à extinção do vínculo do estudante com o seu curso, além de abordar as consequências ocasionadas por esta situação para todos os envolvidos, tais como a instituição de ensino, o próprio estudante, seus familiares e a sociedade em geral. Nesta Seção o construto da evasão escolar é definido sob o ponto de vista de variados autores, conjuntamente com a apresentação de uma análise bibliográfica sobre as causas e consequências relacionadas com a evasão escolar. Apresenta-se, ainda, um breve panorama dos estudos sobre o construto em questão no escopo do IFMG.

O entendimento conceitual sobre o que representa o fenômeno da evasão é fundamental para que as discussões a respeito do tema sejam conduzidas com o devido rigor científico, comumente baseadas apenas na divulgação de aspectos quantitativos que, por conseguinte, relegam ao segundo plano a observação do fenômeno sob um ponto de vista mais amplo e que envolve questões teóricas mais elaboradas ([CASTRO e TEIXEIRA, 2013](#)).

Dada a abrangência de utilização deste construto, muitas vezes os termos “evasão” e “abandono” são empregados como sinônimos, além de outras derivações que buscam, essencialmente, tratar sobre a ruptura do vínculo entre o estudante e seu curso/instituição. Desta forma, quanto mais ampla é a conceituação definida para o fenômeno em si, mais complexa é a compreensão de casos concretos onde não houve a continuidade dos estudos por algum indivíduo ([SILVA FILHO e ARAÚJO, 2017](#)).

No entendimento de [Vitelli e Fritsch \(2016\)](#), diversas perspectivas são associadas com o termo “evasão” nas produções acadêmicas, cristalizando assim a demanda pela construção e definição de um conceito comum capaz de permitir, com certa padronização, o cálculo de sua ocorrência nas mais variadas instituições de ensino. Esta questão é ampliada por [Gomes \(1998\)](#), ao mencionar que existem especificidades nos níveis de ensino que ocasionam diferenças na forma como o conceito de evasão é interpretado, e por [Nagai e Cardoso \(2017\)](#), quando afirmam que a formalização do construto se mostra variável entre as próprias instituições de ensino. Fica latente, portanto, a necessidade de examinar as diversas conceituações empregadas sobre o construto para que seja explicitada a definição utilizada na sequência deste trabalho. Neste sentido, a primeira padronização consiste no uso apenas do termo “evasão” para denotar a ocorrência do fenômeno.

De acordo com [Baggi e Lopes \(2011\)](#), a evasão é caracterizada quando ocorre uma interrupção no ciclo de estudos, qualquer que seja o nível de ensino do estudante. [Fialho e Prestes \(2014\)](#) compartilham definição semelhante, mas advertem que este entendimento

pode ser melhorado quando a parada nos estudos é analisada conjuntamente com respostas que indicam as circunstâncias e os motivos que levaram o estudante a tomar a decisão de não continuar no seu curso. Para as autoras, a redução nas taxas de evasão vai além da conceituação do fenômeno, demandando que as instituições compreendam os diversos fatores que podem ocasionar a ruptura do vínculo em questão.

[Castelo Branco, Nakamura e Jezine \(2017\)](#) entendem a evasão como um ato de escapar, sendo que quando o termo é aplicado no contexto escolar, se refere ao abandono dos estudos. O fenômeno seria caracterizado, portanto, quando o estudante, balizado por uma resolução pessoal ou não, abre mão da sua formação educacional e encerra, mesmo que temporariamente, sua participação nas atividades acadêmicas. [Bueno \(1993\)](#) produz entendimento um pouco diferente quanto à mencionada decisão de abandonar os estudos, afirmando que somente a determinação pessoal do estudante é que a sustenta. No que tange à temporalidade, autores como [José, Broilo e Andreoli \(2011\)](#) e [Matta \(2011\)](#) consideram que a evasão já está materializada inclusive nos casos em que o estudante pode, eventualmente, reverter o seu desligamento, como, por exemplo, nas situações de trancamento de curso. Por outro lado, [Momm e Momm \(2020\)](#) afirmam que a evasão se concretiza somente quando é formalizada a saída permanente do estudante, sem possibilidade de retorno, exceto nos casos de participação em novos processos seletivos; o trancamento do curso e outras situações similares que permitiriam a conversão da saída seriam definidos, de acordo com esta concepção, como uma “interrupção escolar” temporária.

A evasão é caracterizada de modo mais abrangente por [Fritsch, Rocha e Vitelli \(2015\)](#) ao definirem-na como um complexo fenômeno vinculado à desistência dos estudos por qualquer razão. Para esses autores, a evasão consiste em um processo impactado por múltiplas variáveis que atuam interna ou externamente à instituição de ensino, fazendo com que as condições mínimas para a permanência do estudante sejam inviabilizadas. Tal abrangência na definição do construto também pode ser observada no estudo de [Johann \(2012\)](#), onde a evasão é compreendida como o produto do rompimento do vínculo existente até então entre o estudante e o estabelecimento de ensino.

Essa falta de uniformidade na conceituação do construto também havia sido considerada pela [Comissão Especial de Estudos Sobre a Evasão nas Universidades Públicas Brasileiras \(1996\)](#), indicada pelo Ministério da Educação (MEC) para discutir aspectos da evasão nos cursos superiores ofertados pelas IFES. A Comissão optou por conceituar o construto sob três diferentes enfoques, permitindo assim sua aplicação conforme cada objeto de estudo em questão, bem como reduzindo o risco de construir simplificações ou generalizações excessivas. O primeiro enfoque aborda a evasão do curso, identificada quando acontece o desligamento de um estudante do curso superior no qual ele

está regularmente matriculado, independentemente da situação que originou o fato, tais como desistência (definitiva ou temporária), transferência ou exclusão institucional. O segundo enfoque consiste na evasão da instituição, isto é, quando necessariamente o estudante rompe o seu vínculo com a instituição de ensino, não somente com o curso, mas permanece ligado ao sistema de ensino. O último enfoque é a evasão do sistema, constatada a partir da desistência (temporária ou definitiva) do ensino superior pelo estudante. A Comissão optou, em sua pesquisa, por ter como referência a evasão como sendo a desvinculação do estudante do seu curso superior, sem concluí-lo. A justificativa para esta escolha foi a ausência de dados consolidados em âmbito nacional que permitissem a identificação do destino de cada estudante após a sua desvinculação do curso de origem.

No Censo da Educação Superior, publicado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira ([INEP, 2022](#)), é possível verificar a aplicação prática do conceito de evasão no balizamento de análises sobre dados educacionais, em especial aquelas que versam sobre a continuidade dos estudantes em cursos superiores. A evasão ou permanência podem ser identificadas de acordo com três dimensões que, por sua vez, são definidas conforme o vínculo do estudante com a instituição: permanência (quando o vínculo é registrado como “cursando” ou “matrícula trancada”); desistência (o vínculo indica “desvinculado do curso” ou “transferido para outro curso da mesma IES”); conclusão (o estudante já se formou no curso). É possível notar, neste caso, que as análises estatísticas do INEP não consideram como evadidos aqueles estudantes que optaram pelo trancamento do seu curso, fato este que encontra respaldo teórico em alguns autores conforme apresentado anteriormente.

No caso do presente estudo, decidiu-se pela adoção da abordagem apresentada por [Castelo Branco, Nakamura e Jezine \(2017\)](#) no que concerne à compreensão de que a caracterização da evasão pode acontecer a partir do momento em que o estudante não continua sua formação educacional, quaisquer que sejam as razões para tal. Complementarmente, optou-se por respeitar o proposto pelo primeiro enfoque trazido pela [Comissão Especial de Estudos Sobre a Evasão nas Universidades Públicas Brasileiras](#), onde prevalece o entendimento de que a evasão é assinalada a partir do desligamento do estudante do seu curso superior. Considera-se, ainda, a abordagem de [José. Broilo e Andreoli \(2011\)](#) quanto ao aspecto temporal do fenômeno, de maneira que a evasão seja identificada não somente naqueles casos quando o estudante formaliza a ruptura do vínculo, mas também em situações de interrupções temporárias, como trancamento de curso.

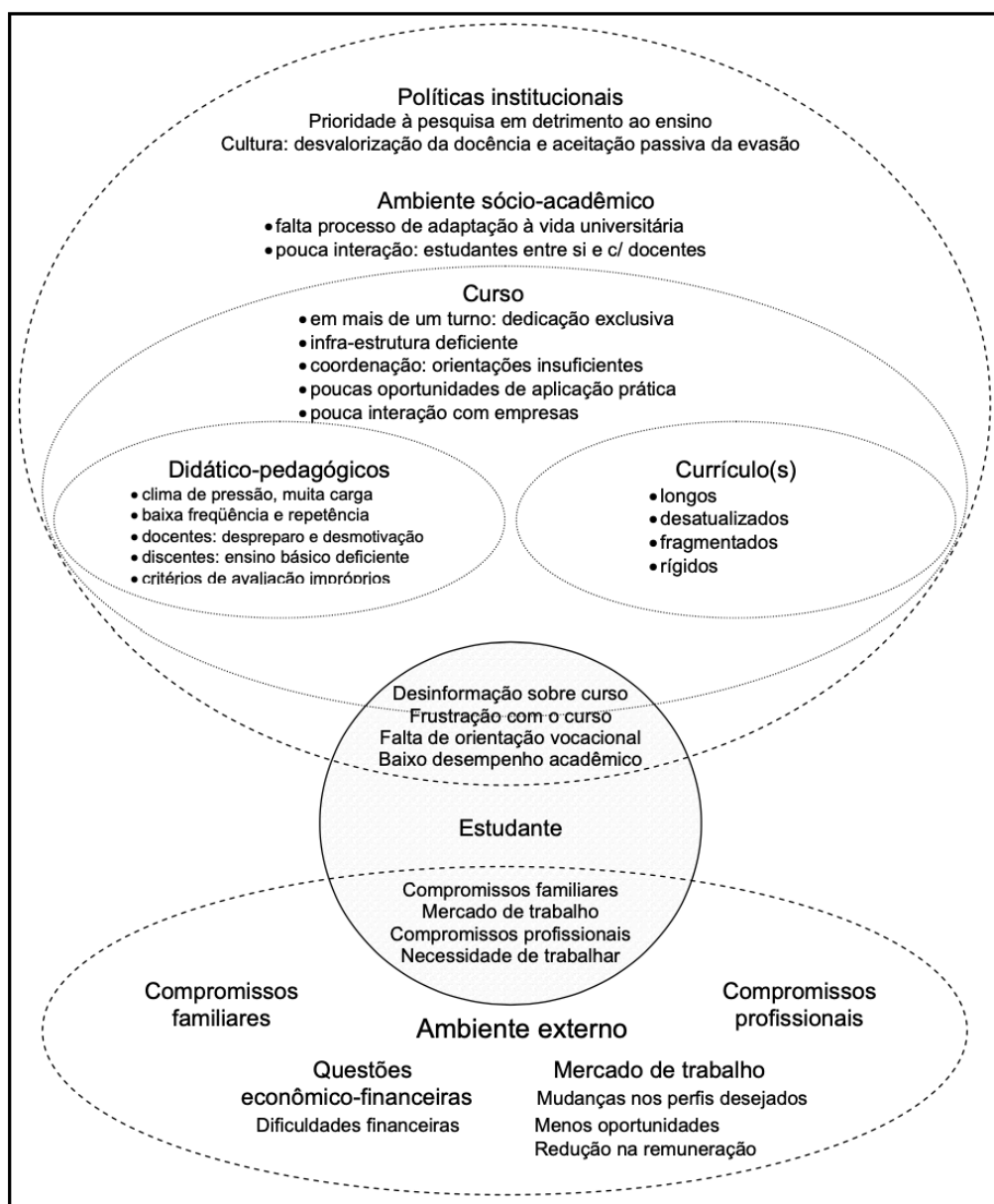
2.1.1 Causas e consequências da evasão

Previamente à apresentação e discussão a respeito das principais causas da evasão escolar no âmbito do ensino superior brasileiro, é válido resgatar a pertinente colocação de Ristoff (1997, *apud* [Coimbra, Silva e Costa, 2019](#)) quanto aos atos de desligamento que invariavelmente acontecem nas instituições de ensino. Para este autor, existe uma quantidade grande de fatores que podem estar ligados à causa da evasão escolar, dentre as quais a própria escolha do estudante pela não continuidade da sua formação acadêmica em uma Instituição de Ensino Superior (IES). Esta situação reflete com certa clareza um ponto que, por vezes, é desconsiderado nas análises envolvendo a evasão escolar: é impossível para qualquer instituição prover continuamente todas as condições para a permanência de seus estudantes. Por outro lado, o conhecimento acerca daqueles fatores que contribuem para a saída dos estudantes é fundamental para que as ações institucionais tenham efeitos positivos naquelas situações onde a ruptura do vínculo ainda pode ser revertida a tempo ([SHIRASU e ARRAES, 2016](#)).

Outro ponto que demanda atenção é mencionado por [Gubert et al. \(2018\)](#) e se refere ao fato de que, normalmente, a evasão não ocorre em virtude apenas de uma causa específica, mas sim de um conjunto de motivos que, conectados, transformam em insuficientes as condições de permanência na instituição. Para os autores, existe certo grau de concordância sobre a evasão ser um fenômeno complexo e impactado paralelamente por mais de uma variável, o que torna a sua compreensão de extrema importância para que os estudantes tenham, na medida do possível, um ambiente capaz de promover aspectos favoráveis à permanência.

[Cislaghi \(2008\)](#) sintetizou visualmente (Figura 2) a grande quantidade de fatores que, eventualmente, estão ligados com a evasão escolar. Nota-se que eles estão relacionados com as mais diversas situações, desde aquelas particulares do estudante, como aspectos familiares e financeiros, até questões localizadas no domínio da própria IES, como a infraestrutura disponível para a realização do curso. O autor ressalta a importância para o controle da evasão do monitoramento contínuo por parte dos gestores institucionais quanto a esses fatores, contrapondo com o entendimento que alguns deles estão fora do alcance de ações ou programas da instituição por serem específicos do estudante.

Figura 2: Fatores que podem influenciar a evasão



Fonte: Cislighi (2008)

De maneira similar à organização proposta por Cislighi (2008), os autores [Biazus \(2004\)](#), [Martins \(2007\)](#) e [Silva Filho et al. \(2007\)](#) classificam as causas da evasão de acordo com sua relação com o ambiente interno ou externo da instituição, além daquelas vinculadas aos aspectos pessoais do estudante. Trata-se da mesma classificação recomendada pela Comissão Especial de Estudos sobre Evasão nas Universidades Públicas Brasileiras e permite observar as causas da evasão de maneira mais estruturada do que caso fossem dispostas individualmente.

As causas internas normalmente envolvem questões didático-pedagógicas (sistema de avaliação, currículo do curso e diversidade de atividades acadêmicas), infraestrutura da

instituição (salas de aula, biblioteca e laboratórios adequados), corpo docente (qualificação, formação e trato com os estudantes) e assistência estudantil (manutenção de programas com apoio socioeconômico suficiente que contribuam para a permanência de estudantes em vulnerabilidade). As questões relacionadas com as causas externas podem advir de aspectos econômicos, políticos e sociais do país, como o mercado de trabalho escasso ou saturado para determinadas áreas profissionais, o custo de vida na cidade onde se localiza a instituição e a logística de transporte necessária para o deslocamento dos estudantes até o campus, dentre outros. As questões pessoais dos estudantes também apresentam bastante variabilidade e, até mesmo, peso proporcional na decisão pela ruptura do vínculo com a instituição, dada a realidade de cada indivíduo. De maneira geral, as causas advêm de assuntos de natureza familiar (gravidez precoce, mudança de cidade/distanciamento da família e problemas de saúde), vocação para a área de estudos (ausência de vínculo com a profissão ou sentimento de inadequação dentro do curso) e, até mesmo, pelo simples interesse de continuar os estudos em nível superior.

Outra categorização das causas da evasão foi introduzida por Schargel e Smink (2002, *apud* [Mello e Santos, 2012](#)), agrupando-as em organizacionais, econômicas, interacionais, psicológicas e sociológicas. As causas organizacionais buscam contribuir para o entendimento do impacto produzido pelas várias características da instituição na decisão do estudante encerrar o vínculo com o seu curso. As questões econômicas envolvem a avaliação, pelo estudante, se existe uma relação positiva entre o que ele financeiramente despende e o retorno obtido com os estudos no nível superior. As causas interacionais dizem respeito à questões de vinculação do estudante com seus pares na instituição, tais como os colegas de classe, funcionários administrativos e docentes. A análise da evasão com base em aspectos sociológicos visa evidenciar que a ruptura do vínculo estudante/instituição não é um ato isolado em si mesmo, mas um elemento que compõe um quadro mais amplo de toda a sociedade. No caso das causas categorizadas como psicológicas, elas revelam aspectos comportamentais e pessoais do estudante como, por exemplo, a imaturidade e rebeldia, tornando-se fatores que contribuem para a evasão.

Considerando os dados apresentados por [INEP \(2022\)](#), o ingresso da maioria dos estudantes de ensino superior no país acontece aos dezenove anos, momento no qual o amadurecimento pessoal talvez ainda não permita uma decisão clara a respeito de qual profissão seguir ([SOUZA e TOASSI, 2014](#)). Paralelamente a este aspecto, há de se considerar que toda escolha, assim como a de um curso superior, envolve um conjunto de expectativas que, por sua vez, podem não ser atendidas na medida esperada pelo estudante, ocasionando frustração com a escolha. Outro fator que contribui indiretamente para que o estudante avalie a possibilidade de interromper o seu vínculo com a instituição é

a opção por um curso baseada nas preferências e pressões familiares, em detrimento de sua própria vontade ([MOMM e MOMM, 2020](#)).

De acordo com [Bueno \(1993\)](#), as possibilidades de ingresso na universidade foram ampliadas a partir de uma conjuntura que envolveu a expansão do ensino superior público brasileiro e a sistematização do processo de distribuição de vagas com o apoio do Sistema de Seleção Unificada (SISU), onde a nota obtida pelo estudante no Exame Nacional do Ensino Médio (ENEM) é utilizada em processos seletivos para instituições de todo o país. Todavia, mesmo tratando de uma inovação que incrementou as condições de mobilidade e acesso dos estudantes, a ruptura parcial dos laços familiares, necessária após eventual mudança de cidade do estudante para o local de oferta do curso escolhido, pode contribuir para os índices de evasão na medida em que ocasiona maiores dificuldades para aqueles que não possuem condições adequadas para se adaptar a esta nova realidade ou não estejam pessoalmente preparados para tamanha transformação.

Aspectos sobre a relação entre o estudo e a necessidade de trabalhar para garantir o sustento são analisadas por [Silva Filho e Araújo \(2017\)](#). Considerando a ampla desigualdade existente no Brasil, em diversos sentidos, é plausível conceber que uma mesma turma abarque desde estudantes com condições de se dedicarem integralmente ao curso até aqueles que necessitam trabalhar durante meio período ou todo o dia de forma a custear financeiramente os gastos relacionados com a realização de um curso superior, como, por exemplo, deslocamento, alimentação e aquisição de materiais. A partir do momento em que o estudante percebe que não está sendo viável o compartilhamento do seu tempo entre as atividades laborais e de estudo, bem como os gastos para sua manutenção no curso são superiores aos rendimentos auferidos pelo trabalho, ele pode entender que o encerramento do seu curso, no momento, é a opção mais adequada para sua realidade.

[Saccaro, França e Jacinto \(2016\)](#) e [Silva e Costa \(2018\)](#) abordam como as restrições na distribuição de auxílios socioeconômicos contribuem para a evasão de estudantes matriculados em cursos superiores. Conforme esses autores, o contexto trazido pelo Reuni, especialmente no que tange à adequação da infraestrutura das IES para disponibilizar um número maior de vagas e cursos, fomentaram indiretamente os programas de assistência estudantil, dada a necessidade de apoiar maior quantitativo de estudantes com comprovada dificuldade de manutenção do vínculo com o curso. Contudo, o aumento no número de matrículas no ensino superior brasileiro não vem sendo acompanhado, proporcionalmente, pelo volume de recursos financeiros atribuídos às instituições via PNAES, fazendo com que a falta deste suporte à permanência estudantil colabore para o aumento da evasão.

Outro fator que pode impactar negativamente na evasão se refere ao despreparo ou baixa qualificação do corpo docente, conforme abordado por [Santos et al. \(2019\)](#) e [Silva e](#)

[Ramos \(2018\)](#). Entende-se que este fator pode, por outro lado, contribuir para a permanência dos estudantes, na medida em que o docente com adequada capacidade de atuação e relacionamento consegue motivar sua turma a continuar os estudos ([SILVA, 2019](#)). Entretanto, caso o estudante perceba, mesmo que involuntariamente, que o seu conjunto de professores não está apto a entregar tudo aquilo que é esperado em termos de aprendizado ou condução das aulas, existe a possibilidade dele avaliar se o esforço para continuar estudando na instituição e no curso são compensatórios ou existem perspectivas de mudanças a curto prazo, por exemplo, em função da atuação de novos docentes no próximo semestre de estudos. No caso daquelas instituições que não promovem um acompanhamento permanente do desempenho dos docentes junto às suas turmas, esta situação pode se agravar e tomar maiores dimensões.

Para [Figueiredo e Salles \(2017\)](#), a evasão escolar também sofre influência da defasagem de conhecimentos prévios dos estudantes, sendo este um fator que demanda monitoramento contínuo por parte dos gestores acadêmicos. Atualmente, o cenário da educação superior no Brasil é composto por estudantes que cursaram o Ensino Médio, na sua maioria, em escolas públicas ([ANDIFES, 2019](#)), onde infelizmente a qualidade da aprendizagem é deficitária, mesmo quanto aos conhecimentos básicos de escrita e leitura. Portanto, a realidade aponta para uma alta probabilidade de ingresso nas IES de estudantes com defasagens de conhecimentos prévios. Analisando apenas as disciplinas do primeiro ano dos cursos de graduação, notadamente aquelas ofertadas nos bacharelados, os conteúdos prévios adquiridos ao longo do Ensino Médio são fundamentais para que os estudantes consigam acompanhar as demandas específicas do seu curso. A evasão pode acontecer a partir da percepção, pelo estudante, de que o seu entendimento sobre determinados conteúdos não é satisfatório à obtenção das condições mínimas para aprovação, mesmo aplicando certo esforço no processo de aprendizado. Ademais, consonante com o exposto por [Sadoyama et al. \(2020\)](#) e [Veloso, Couto e Valentim \(2018\)](#), há de se ressaltar que nem todo docente e a própria instituição observam com a desejada frequência tal situação em seus estudantes, ou quando o fazem, assumem uma postura passiva no sentido de somente aguardar a evolução do estudante sem oferecer os meios para que isto seja possível, como estudos de reforço ou nivelamento.

Ao abordarem as consequências geradas pela evasão escolar, [Nagai e Cardoso \(2017\)](#) afirmam que a interrupção de um curso eventualmente implica em dificuldades pessoais do estudante, como sentimentos de incapacidade intelectual, fracasso e frustração, problemas materiais, como dívidas contraídas para a permanência e participação no curso, e dificuldades profissionais em razão da ausência de qualificação acadêmica formal. Entendimento similar possuem [Castro e Teixeira \(2013\)](#), indicando que essa ruptura nos estudos pode ocasionar, além de prejuízos financeiros extensíveis à família do estudante,

possíveis problemas psicológicos nos casos em que a situação não seja bem compreendida e trabalhada pelos indivíduos. No que tange aos aspectos da exclusão social como consequência da evasão, [Ceratti \(2008\)](#) entende que este é um fenômeno que deve ser prontamente enfrentado pelas IES, pois o papel destas perante à sociedade é justamente oferecer um ambiente de estudos que contemple inúmeras possibilidades até então fora da realidade de muitos estudantes com dificuldades socioeconômicas.

De acordo com [Silva Filho et al. \(2007\)](#), os prejuízos advindos da evasão escolar provocam impactos diretos não somente nas instituições, mas em todos os componentes do sistema educacional. Tanto no setor privado quanto no público, o volume de recursos financeiros, estruturais e humanos desperdiçados é uma realidade que, por si só, já demanda a condução de investigações contínuas que permitam observar quais práticas surtem efeitos benéficos capazes de reduzir a ocorrência da evasão. No âmbito das instituições públicas, [Cislaghi \(2008\)](#) ressalta que, se os principais fatores envolvidos com o fenômeno da evasão fossem ao menos identificados e acompanhados por meio de programas institucionais, seria possível controlar o desperdício de verbas acumulado ao longo de todas as interrupções de vínculos estudantis. Para este autor, o contexto da evasão nas IFES é ainda mais preocupante, pois a maior fatia do orçamento governamental direcionado a elas advém dos resultados da permanência e da evasão. O estudo de [Gonçalves \(2018\)](#) dimensiona quantitativamente o tamanho deste impacto nas IFES, demonstrando como a Universidade Federal de Alfenas (UNIFAL-MG) deixou de receber, entre 2015 e 2018, mais de 71 milhões de reais via matriz orçamentária, dados os índices de evasão verificados em seus cursos superiores.

A evasão também gera consequências relativas à colocação profissional, pois conforme abordam [Linke, Nogueira e Linke \(2017\)](#), estudantes podem ter dificuldades de empregabilidade no contexto de um mercado de trabalho marcado pela exigência de formação em nível superior como requisito mínimo para admissão. A própria manutenção no emprego, em diversos casos, passa a estar condicionada a um vínculo do profissional com algum curso superior. Para [Alves \(2008\)](#), o alcance do impacto da evasão não pode ser analisado apenas sob o ponto de vista dos estudantes universitários, mas precisa envolver as empresas que buscam novos talentos para recrutamento, pois a partir do momento em que os índices de evasão estão altos, reduz-se a quantidade de potenciais profissionais qualificados que poderiam ser contratados.

[Davok e Bernard \(2016\)](#) propõem que uma das consequências de estudos sobre evasão, dentro de uma abordagem organizacional, consiste na indicação formal para os gestores acadêmicos sobre qual é a caracterização do fenômeno dentro da IES, esta construída com o suporte de técnicas sistematizadas para a coleta e o uso dos dados educacionais. Não obstante, o emprego dos resultados deste estudo para compreensão do

fenômeno em si, é fundamental que a instituição também utilize-o para o desenvolvimento de políticas e programas institucionais que contribuam com a manutenção do vínculo dos estudantes com seus cursos.

2.1.2 A evasão no IFMG

O IFMG divulgou, em 2017, um documento intitulado “A evasão escolar no IFMG: diagnóstico e diretrizes da política institucional para a permanência e êxito dos estudantes”, visando apresentar o panorama geral da evasão em seus cursos com base em dados obtidos em sistemas informatizados dos ingressantes entre 2008 a 2015 e dados compilados após aplicação de questionários aos estudantes evadidos no ano de 2015. O conjunto de análises compreendeu a caracterização socioeconômica dos estudantes, a diferenciação entre o perfil dos estudantes que continuaram matriculados e daqueles que evadiram, bem como a construção de considerações sobre os fatores citados como origem da evasão. As atividades foram conduzidas por uma comissão designada com o intuito de elaborar as linhas gerais da política institucional de auxílio e permanência dos estudantes, envolvendo cursos de todas as modalidades ofertadas pela instituição (IFMG, 2017).

Quanto aos concluintes e evadidos, os resultados da instituição eram preocupantes, conforme observa-se na Tabela 1. Considerando cada modalidade de curso ofertada, a evasão média era de 28,2% nos cursos de Bacharelado, 36,6% nos cursos Tecnólogos e 44,8% nas Licenciaturas. Apesar de ter grande variação entre as modalidades, as taxas médias dos concluintes eram baixas, não ultrapassando 15%.

Tabela 1: Situação dos estudantes de cursos superiores do IFMG (2008 a 2015)

Indicador	Bacharelado	Tecnólogo	Licenciatura
Estudantes	3.935	1.611	1.356
Evadidos	1.110 (28,2%)	590 (36,6%)	382 (28,2%)
Matriculados	2.589 (65,8%)	788 (48,9%)	892 (65,8%)
Concluintes	236 (6%)	233 (14,5%)	82 (9%)

Fonte: Adaptado de IFMG (2017)

Como o documento do IFMG apresenta resultados gerais que envolvem não apenas cursos de graduação, torna-se impraticável construir diferentes caracterizações do perfil dos estudantes analisados. Uma das únicas informações segmentadas pela modalidade é o percentual de evasões registradas no primeiro ano letivo do curso. Nos Bacharelados, 41% das evasões ocorreram no primeiro período e 24% no segundo período, totalizando 65% no primeiro ano letivo. Nas Licenciaturas, o resultado final do ano foi de 69%, sendo 50%

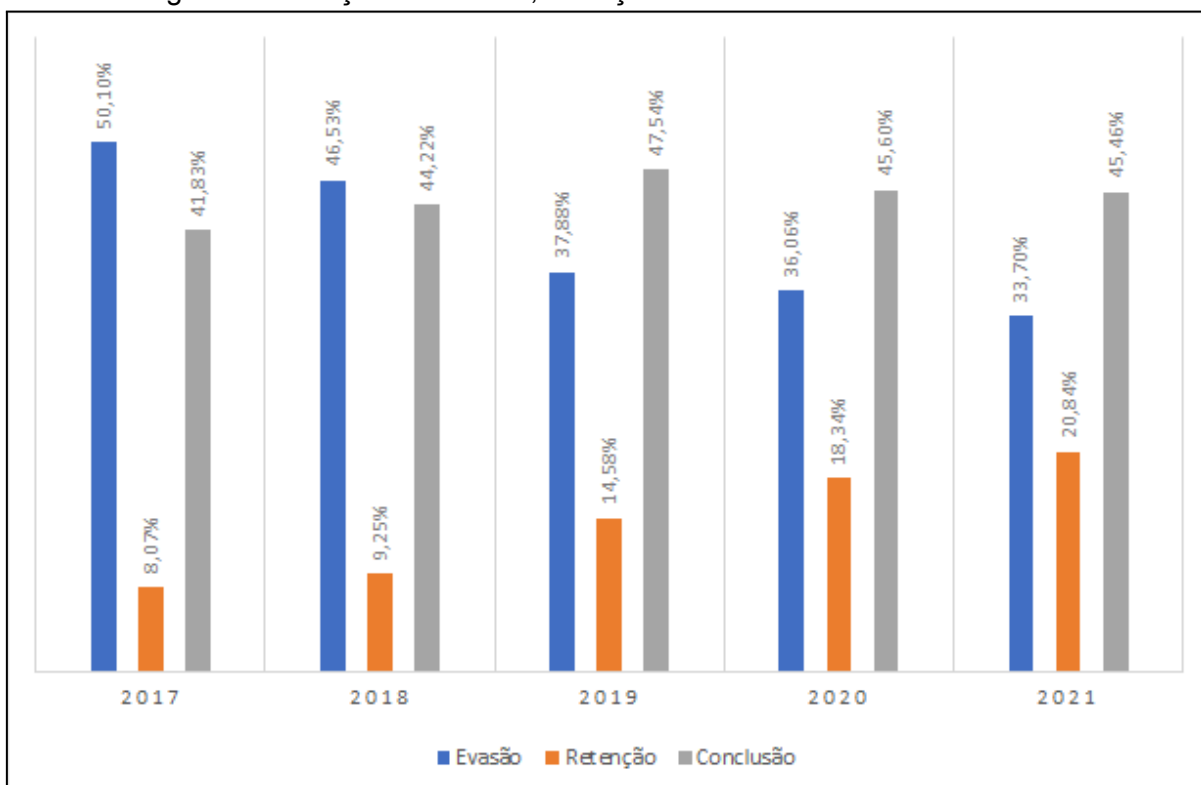
durante o primeiro período e 19% no segundo. O maior percentual total estava relacionado com os Tecnólogos (70%), com 46% no primeiro período e 24% no segundo.

A pesquisa do IFMG coletou junto aos alunos evadidos em 2015 quais foram os principais fatores (internos e externos) que contribuíram para o encerramento do vínculo com o curso. Dentre os fatores externos mais mencionados pelos participantes, destacam-se as dificuldades financeiras para se manter estudando, a pequena identificação com o curso, problemas para acessar o Campus, rendimento insuficiente em determinados componentes curriculares e percalços para conciliar o trabalho com a participação nas atividades acadêmicas. Quanto aos fatores internos, foram citadas a alta carga horária dos cursos, a falta de ambientes adequados para aprendizado de disciplinas práticas e a ausência de suporte adequado aos docentes.

No documento consta, ainda, uma seção onde são propostas doze ações para auxiliar o êxito e a permanência dos estudantes do IFMG. Entretanto, nas consultas realizadas no *website* da instituição, não foi encontrada nenhuma informação indicando o acolhimento das ações como integrante de algum tipo de política institucionalizada para gestão da evasão. Observando o relatório de gestão publicado pelo IFMG em 2022, o termo “evasão” sequer foi mencionado ([IFMG, 2022a](#)).

Analisando os resultados da evasão, retenção e conclusão nos cursos do IFMG disponibilizados pela plataforma Nilo Peçanha ([BRASIL, 2022](#)), os quais referem-se ao ciclo de cada turma, isto é, o período compreendido entre o semestre de ingresso no curso e o semestre de encerramento conforme o prazo para integralização mínima, verificam-se números animadores quanto à evasão e outros merecedores de atenção relativamente à retenção (quando o estudante continua matriculado para cumprir créditos acadêmicos mesmo quando ultrapassado o tempo mínimo previsto para integralização do seu curso). É possível visualizar, na Figura 3, a queda consistente no percentual de evasão entre 2017 e 2021, alterando o contexto no qual praticamente a metade dos estudantes que ingressavam em determinado curso evadiram. Por outro lado, como as taxas de conclusão se mantiveram estáveis ao longo do período analisado, é possível observar a evolução negativa do indicador de retenção, permitindo afirmar que mais estudantes estão tendo dificuldades para concluir seus cursos.

Figura 3: Evolução da evasão, retenção e conclusão entre 2017 e 2021



Fonte: Elaborado a partir dos dados coletados em [BRASIL, 2022](#)

A efetividade das ações de combate à evasão escolar nas unidades da Rede Federal de Educação Profissional, Científica e Tecnológica (EPCT), da qual o IFMG faz parte, vem sendo objeto de auditorias executadas pelo Tribunal de Contas da União (TCU). Dentre os fatores relacionados negativamente com a evasão, o TCU destacou que os recursos financeiros destinados às instituições são insuficientes para atender a demanda das ações de assistência estudantil e alimentação escolar, além do número reduzido de servidores para composição de equipes multiprofissionais para tratarem desse assunto sob o ponto de vista organizacional. Ademais, o órgão concluiu que, de maneira geral, o planejamento estratégico das instituições para fomentar a permanência e o êxito dos estudantes em seus cursos continha deficiências que contribuam para o quadro de alta evasão ([TCU, 2025](#)).

Nesse sentido, em 2024 o IFMG formalizou o seu Plano Institucional de Permanência², composto de um conjunto de ações e metas a serem conduzidas por cada campus objetivando reduzir a evasão escolar e oferecer condições adequadas para a manutenção do vínculo dos estudantes até o encerramento com sucesso dos seus cursos. A execução das ações se dará entre 2025 e 2027, sendo monitoradas pela Reitoria com o apoio de comissões locais nos *campi*.

²

https://www.ifmg.edu.br/porta/ensino/normas-internas-1/Plano_Institucional_de_Permanencia_IFMG_.pdf

Alguns estudos acadêmicos analisaram o fenômeno da evasão no âmbito do IFMG, porém considerando apenas um curso ou Campus e não toda a instituição. [Castro \(2019\)](#) analisou a evasão em cursos de Licenciatura ofertados pelos *campi* Bambuí, Congonhas e Ouro Branco, objetivando identificar quais fatores estavam associados com o processo de desligamento. Apesar do recorte limitado em relação à variedade e quantidade de cursos ofertados pelo IFMG, a pesquisa encontrou evidências de que, nos casos estudados, a evasão estava relacionada com múltiplos fatores pessoais, sociais e até mesmo da carreira docente. Outra importante ponderação da autora indica a ausência de programas institucionais voltados para a contenção da evasão no âmbito do IFMG e, por conseguinte, melhorias nos índices de permanência.

[Conceição, Longhini e Oliveira \(2020\)](#) identificaram o perfil dos 420 estudantes evadidos no curso de Engenharia de Produção do Campus Governador Valadares entre 2010 e 2019. Observou-se que a maioria dos estudantes possuía menos de 22 anos, efetivou a interrupção dos estudos até o terceiro período do curso e possuía renda *per capita* familiar de até um salário mínimo. Os autores ressaltam que o Campus não executa nenhum tipo de ação preventiva à evasão, tais como avaliação periódica da frequência dos estudantes ou acompanhamento da renovação de matrícula.

[Assunção \(2021\)](#) conduziu uma pesquisa com 830 estudantes do Campus Ouro Preto para compreender quais fatores contribuem para a permanência na instituição e quais estão mais vinculados com a evasão. No caso da permanência, destaca-se a oferta de condições pela instituição para que os estudantes possam desempenhar o papel que deles é esperado, tais como ambientes com infraestrutura adequada e organização didático-pedagógica. Quanto à evasão, alguns pontos podem ser melhorados, como maior apoio socioeconômico aos estudantes com vulnerabilidades e incremento na oferta de atividades diversificadas em contraponto à tradicional aula em sala.

2.2 Assistência estudantil

Nesta Seção é feita uma breve conceituação e contextualização histórica da assistência estudantil no âmbito das IFES, de maneira a destacar o seu papel como suporte à manutenção do vínculo estudantil para aqueles com vulnerabilidades comprovadas. Apresenta-se, ainda, a configuração atual da assistência estudantil no IFMG e como ocorre o processo de seleção para distribuição dos auxílios socioeconômicos, permitindo compreender como esta tese pode contribuir para a melhoria institucional no que tange ao

entendimento sobre as características dos estudantes que demandam tais auxílios e qual o perfil de comportamento informacional dos gestores desta área.

Em termos conceituais, [Martins, Silva e Maurício \(2019\)](#) definem que a assistência estudantil é uma forma de direito social dos indivíduos, visando fornecer recursos para que os estudantes superem obstáculos, se desenvolvam como pessoas e como profissionais ao longo do curso e consigam adequado desempenho acadêmico, reduzindo, por outro lado, o volume de trancamentos e abandonos de matrículas. Para uma das principais referências deste assunto no Brasil, o Fórum Nacional de Pró-Reitores de Assuntos Comunitários e Estudantis (FONAPRACE), a assistência estudantil pode ser entendida como uma política composta por uma série de diretrizes e princípios, baseados na perspectiva da inclusão social e da formação ampliada, cujo objetivo é guiar as IFES na elaboração de ações que, por sua vez, garantam o acesso, a permanência e a conclusão dos cursos de estudantes ([FONAPRACE, 2012](#)).

Para [Barbosa \(2009\)](#), a assistência estudantil é qualificada como responsável pela proposição e execução de projetos que auxiliem diretamente os estudantes a permanecerem em seus cursos e estabelecimentos de ensino, promovendo aos beneficiados melhor qualidade de vida e, conseqüentemente, meios adequados para explorar todo o potencial de desempenho acadêmico de cada indivíduo. Conforme [Back \(2019\)](#), é fundamental analisar a assistência estudantil inserida no contexto da própria política de educação superior do país, pois sua concepção não pode ser dissociada do processo de garantia do direito à educação, sendo esta uma das suas principais contribuições. [Silva e Nogueira \(2015\)](#) afirmam que a assistência estudantil, dada sua pertinência, é uma condição estruturante para a permanência dos estudantes no ensino superior.

De acordo com [Vasconcelos \(2010\)](#), as ações da assistência estudantil se materializam na forma de programas e projetos institucionais focados no aspecto educacional, diferenciando-se, portanto, de práticas caracterizadas como caridade. Tais ações normalmente estão relacionadas a questões de saúde dos estudantes, apoio pedagógico para necessidades educativas específicas, além do provimento dos meios à sobrevivência do estudante, englobando alimentação, transporte, moradia e recursos financeiros. Ao tratarem sobre o escopo destas ações, [Finatti, Alves e Silveira \(2007\)](#) afirmam que o adequado desenvolvimento acadêmico dos estudantes somente será alcançado quando houver, paralelamente, uma associação entre a qualidade do ensino ofertado com um efetivo conjunto de ações da assistência estudantil, sendo este baseado no contexto experimentado pela instituição. [Pereira e Almeida \(2015\)](#) ressaltam o grande desafio que é a elaboração de projetos capazes de, ao mesmo tempo, equilibrar aspectos do orçamento financeiro disponível e atender as condições mínimas para permanência e

apoio constante a um número cada vez maior de estudantes com vulnerabilidades e/ou demandas específicas.

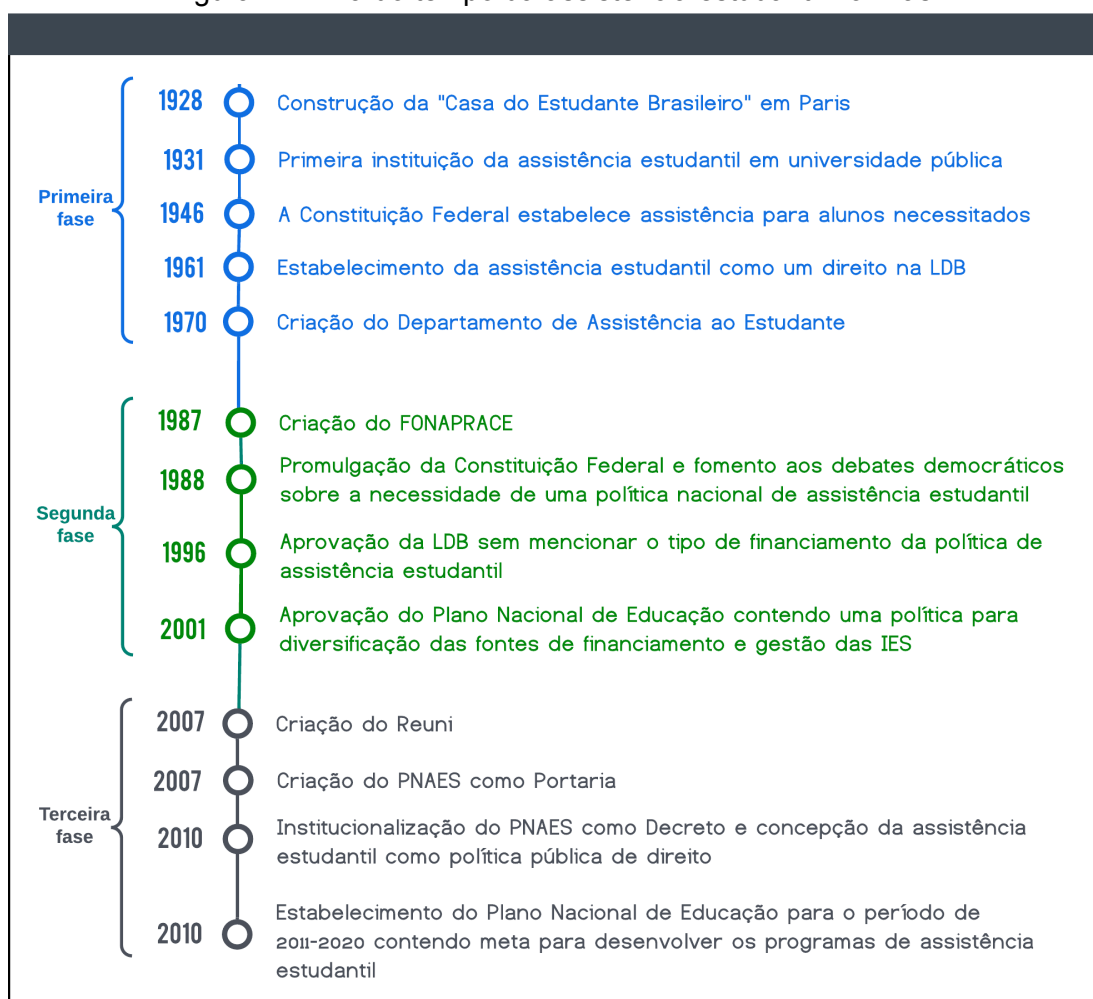
[Barbosa \(2009\)](#) e [Santos e Freitas \(2014\)](#) fazem uma crítica no sentido de ampliar o entendimento da assistência estudantil como direito social, de maneira que o seu caráter focalizado e seletivo, muito em função de restrições orçamentárias e de pessoal, dê espaço para um formato que promova efetivamente a universalidade do acesso e permanência no ensino superior. [Pereira e Stein \(2010\)](#) propõem raciocínio semelhante ao afirmarem ser preciso que a assistência estudantil não funcione como uma política pública segmentada e referenciada exclusivamente pela renda, pois o seu escopo de atuação deve permitir a amplitude da assistência em um universalismo real e não contido em si mesmo.

Ainda neste sentido de crítica à estruturação da assistência estudantil, [Cislaghi e Silva \(2012\)](#) ressaltam a necessidade de distinção entre esta e a política de assistência social, para que sua atuação ocorra dentro da fronteira do direito à educação como política universal, e não apenas replicando a lógica da assistência social de focalização e bolsificação. Para [Angelim \(2010\)](#), é preciso encontrar formas de fazer com que a assistência estudantil não seja focalizada e baseada simplesmente na concessão de auxílio financeiro para apoio aos estudantes, mas que prevaleça a oferta de um agrupamento de serviços de assistência derivado da real necessidade dos indivíduos.

A história da assistência estudantil em muito se confunde com a própria história das Instituições Federais de Ensino Superior brasileiras, dada sua evolução como fonte de apoio estudantil ter ocorrido neste âmbito. [Santos e Marafon \(2016\)](#) mencionam que ações de assistência estudantil são conduzidas em universidades brasileiras desde a década de 1920, mesmo sendo de caráter pontual e abrangência fragmentada. [Imperatori \(2017\)](#) promove um amplo panorama da trajetória da assistência estudantil, definindo como marco inicial de ações de apoio a inauguração, em 1928, da Casa do Estudante Brasileiro em Paris, cuja proposta era auxiliar estudantes universitários com dificuldades para se manter na capital francesa. Inserido neste contexto, a autora apresenta outros momentos históricos que merecem menção por terem contribuído para a evolução da assistência estudantil: as políticas educacionais do governo Getúlio Vargas, a Lei de Diretrizes e Bases da Educação Nacional (LDB) de 1961, a primeira fase de expansão da educação superior entre 1950 e 1970, a criação do Departamento de Assistência ao Estudante (DAE) no âmbito do Ministério da Educação e Cultura, a Constituição Federal de 1988 e a LDB de 1996. Ressalta-se que, apesar de inúmeras tentativas de formalizar a atuação da assistência estudantil sob um entendimento comum, a realidade das instituições de ensino somente permitia a consecução de ações pulverizadas e, na maioria das vezes, baseadas em relações clientelistas ([BARBOSA, 2009](#)).

O resgate histórico promovido por [Kowalski \(2012\)](#) também se mostra muito abrangente e detalhado, abarcando desde os primórdios da educação superior no Brasil e a concepção de políticas públicas educacionais até a evolução da assistência estudantil. Não obstante o fornecimento de criterioso histórico, a autora propõe o agrupamento dos marcos da assistência estudantil em três fases. A primeira fase engloba o maior período, qual seja, partindo da criação da Casa do Estudante Brasileiro até próximo ao momento de democratização política do país em 1988. A segunda fase possui vínculo com um período da história brasileira no qual foram conduzidos vários debates democráticos e, por conseguinte, elaborados novos projetos na área de educação que geraram uma identidade particular da assistência estudantil nas IFES. A última fase teve início com o período de reestruturação e expansão das IFES promovido pelo Governo Federal em 2007, culminando na institucionalização de uma política pública para a assistência estudantil em 2010. A Figura 4 relaciona, no formato de linha do tempo, alguns dos principais marcos citados pela autora dentro do contexto desta evolução no Brasil.

Figura 4: Linha do tempo da assistência estudantil no Brasil



Fonte: Adaptado de Kowalski (2012)

Considerando o escopo desta tese e as três fases delineadas por [Kowalski \(2012\)](#), optou-se por discutir, na sequência, acontecimentos a partir de 2007 que contribuíram diretamente para mudanças na configuração da assistência estudantil nas IFES. O primeiro fato marcante foi a criação do Programa de Reestruturação e Expansão das Universidades Federais (Reuni), a partir da publicação do Decreto Nº 6.096 de 24 de abril de 2007 ([BRASIL, 2007a](#)), composto de várias metas que, uma vez cumpridas, permitiriam expandir o acesso e a permanência no ensino superior por meio de melhorias na infraestrutura física das IFES, contratação de novos servidores, abertura de novos *campi* e cursos, aumento de vagas e oferta de turmas no período noturno ([KOWALSKI, 2012](#)). [Maciel \(2014\)](#) apresenta em sua pesquisa um amplo detalhamento do programa, desde os aspectos teóricos até questões de sua implementação nas universidades participantes. Enquanto para alguns autores, como [Aranha, Pena e Ribeiro \(2012\)](#), [Bittencourt, Ferreira e Brito \(2017\)](#), [Melo et al. \(2020\)](#), [Paula e Almeida \(2020\)](#) e [Santos \(2022\)](#), o Reuni conseguiu alcançar seu propósito e viabilizar o ingresso e a permanência de estudantes com os mais variados perfis no ensino superior, bem como ampliar a oferta de cursos para o interior do país, para autores como [Cañaverl e Sá \(2017\)](#), [Oliveira et al. \(2014\)](#), [Paula \(2017\)](#) e [Silva e Borowski \(2018\)](#), o Reuni não entregou os resultados previstos em termos de eficiência acadêmica e adequado suporte ao trabalho docente.

O segundo fato que demanda menção é a promulgação da Lei Nº 12.711, de 29 de agosto de 2012 ([BRASIL, 2012b](#)), conhecida como Lei das Cotas, cujo objetivo era ampliar ações afirmativas ao garantir um percentual mínimo de 50% das vagas nas IFES para os candidatos que cursaram todo o ensino médio em escolas públicas. A metade das vagas reservadas deveria ser destinada a candidatos com renda *per capita* familiar igual ou inferior a um salário mínimo e meio, aumentando, portanto, as possibilidades de estudantes de baixa renda ingressarem no ensino superior. Para [Senkevics e Mello \(2019\)](#), a partir do momento em que o acesso ao ensino superior público foi democratizado e aspectos de inclusão e diversidade passaram a ser considerados nos processos seletivos, como na Lei das Cotas, entre outras políticas públicas, houve considerável variação no perfil dos ingressantes. Em termos quantitativos, ao final de 2015 cerca de 77% deles eram de baixa renda (estavam dentro da faixa mínima de renda estipulada na reserva de vagas) e 64% estudaram somente em escolas públicas durante o ensino médio. [Trevisol e Nierotka \(2015\)](#) ressaltam o caráter padronizador da Lei das Cotas para todas as IFES, dado que antes do seu efetivo cumprimento, pelo fato de cada instituição adotar parâmetros próprios, o acesso ao ensino superior nem sempre estava ao alcance da população marginalizada acadêmica e economicamente.

É válido ressaltar que a democratização do ensino superior não deve ser medida apenas por melhorias no acesso dos estudantes, mas também pelas condições para que

estes permaneçam na instituição e consigam obter êxito em seus estudos no tempo adequado ([ZAGO, 2006](#)). No entendimento de [Back \(2019\)](#), a oferta de maior quantidade de vagas nas IFES não é, por si só, fator suficiente para que os estudantes consigam concluir o curso com desenvolvimento acadêmico satisfatório ou, até mesmo, tenham condições de permanecer vinculados à instituição, em especial quando considerados os percentuais de ingressantes oriundos de famílias com vulnerabilidade socioeconômica. [Dias e Costa \(2016\)](#) complementam este panorama, indicando que os desafios quanto à permanência no curso enfrentados pelos ingressantes em uma IFES envolvem desde a necessidade de obter recursos financeiros para sobrevivência (moradia, alimentação, vestuário) e para o estudo (livros, materiais e transporte), quanto resiliência para enfrentar dificuldades de cunho acadêmico ou pessoal.

A preocupação com formas de incentivar a permanência no ensino superior, para além da questão da democratização do acesso, também foi objeto de estudo por [Santos e Freitas \(2014\)](#), que apontam as políticas assistenciais como elemento central para a manutenção de estudantes advindos de segmentos sociais historicamente excluídos em termos sociais e econômicos. [Alves \(2002\)](#) entende que o adequado desenvolvimento acadêmico dos estudantes e a sua permanência no ensino superior estão associados não somente com a qualidade de ensino ofertada pela instituição, mas em grande parte a uma efetiva política de assistência que abranja necessidades como a alimentação, transporte, moradia, saúde e lazer, dentre outras condições. Diante desta constatação, a autora defende que é papel das IFES promover a assistência estudantil como um direito de seus estudantes, propondo e executando ações que abarquem integralmente o sujeito em toda a sua formação, indo além, portanto, da mera distribuição de recursos financeiros para aqueles que possuem vulnerabilidades socioeconômicas.

Por outro lado, [Silva e Nogueira \(2015\)](#) apresentam ressalvas quanto à expectativa de associação incondicional entre permanência e assistência estudantil, ponderando que a manutenção do vínculo do estudante com a instituição de ensino não está baseado apenas no atendimento momentâneo à carência material ou financeira do mesmo. Assim como na evasão, um conjunto de fatores atua (muitas vezes de maneira interligada) para que as condições de permanência do estudante sejam favoráveis, dentre as quais é possível relacionar, não exclusivamente, o apoio da assistência estudantil. Para os autores, a permanência vincula-se a fatores cambiáveis com o tempo e podem envolver os desejos particulares do estudante, sua rede de relacionamentos na instituição, as características do ambiente acadêmico, além de situações que fogem do alcance gerencial, como problemas familiares que impactam diretamente na base que mantém o vínculo do estudante com o seu curso. [Back \(2019\)](#) salienta que, mesmo a assistência estudantil não sendo capaz de, por si só, garantir a permanência dos estudantes, a sua importância se dá justamente pela

capacidade de acompanhá-los ao longo da sua vida acadêmica e identificar aqueles momentos nos quais as condições mudaram e se apresentam os fatores que tornam complexa a manutenção do relacionamento estudantil.

Neste contexto de ampliação das possibilidades de acesso ao ensino superior e demanda por meios para apoiar a permanência dos ingressantes em seus estudos, no final de 2007 foi criado o PNAES como uma proposta de formalizar o atendimento às demandas acadêmicas e pessoais de estudantes formalmente vinculados a cursos de graduação no âmbito das IFES. Em termos legais, sua instituição ocorreu por meio da Portaria Nº 39, de 12 de dezembro de 2007 ([BRASIL, 2007b](#)), sendo devidamente regulamentado pelo Decreto Nº 7.234, de 19 de julho de 2010 ([BRASIL, 2010](#)).

Embora as IFES já prestassem assistência aos estudantes, a formalização do PNAES serviu como um instrumento norteador para o planejamento de ações que buscassem aumentar, prioritariamente, as condições positivas para a permanência dos estudantes de cursos superiores ([SANTOS e FREITAS, 2014](#)). [Kowalski \(2012\)](#) ressalta que, apesar de o PNAES não se constituir como mecanismo exclusivo para fomentar a permanência daqueles estudantes com vulnerabilidade socioeconômica identificada, sua formalização pode ser compreendida como um marco na regulação das políticas de assistência, fazendo com que a própria estrutura das IFES passasse por modificações que espelhavam a nova importância deste setor dentro da organização institucional.

No entendimento de [Santos e Marafon \(2016\)](#), a formalização do PNAES como proposta governamental permitiu que sua abrangência fosse ampliada para ações que não devem se basear somente na distribuição de auxílios monetários, mas sobretudo no acompanhamento de diversos aspectos da vida pessoal e acadêmica dos estudantes. Neste sentido, o Decreto especifica dez áreas nas quais as ações da assistência estudantil devem ser desenvolvidas: moradia estudantil, alimentação, transporte, atenção à saúde, inclusão digital, cultura, esporte, creche, apoio pedagógico e acesso de estudantes com demandas específicas. Para [Santos \(2020\)](#), o delineamento destas áreas serviu de modelo para que as IFES desenhassem, com certo grau de padronização, suas políticas institucionais de assistência estudantil. [Gomes e Passos \(2018\)](#) complementam que o PNAES foi definido com cinco objetivos que, analisados sob o ponto de vista integrado, norteiam a elaboração de ações para contribuir com a permanência dos estudantes, redução da evasão escolar, melhoria do desempenho acadêmico, promoção da inclusão social e redução do impacto das desigualdades sociais e regionais na conclusão dos estudos.

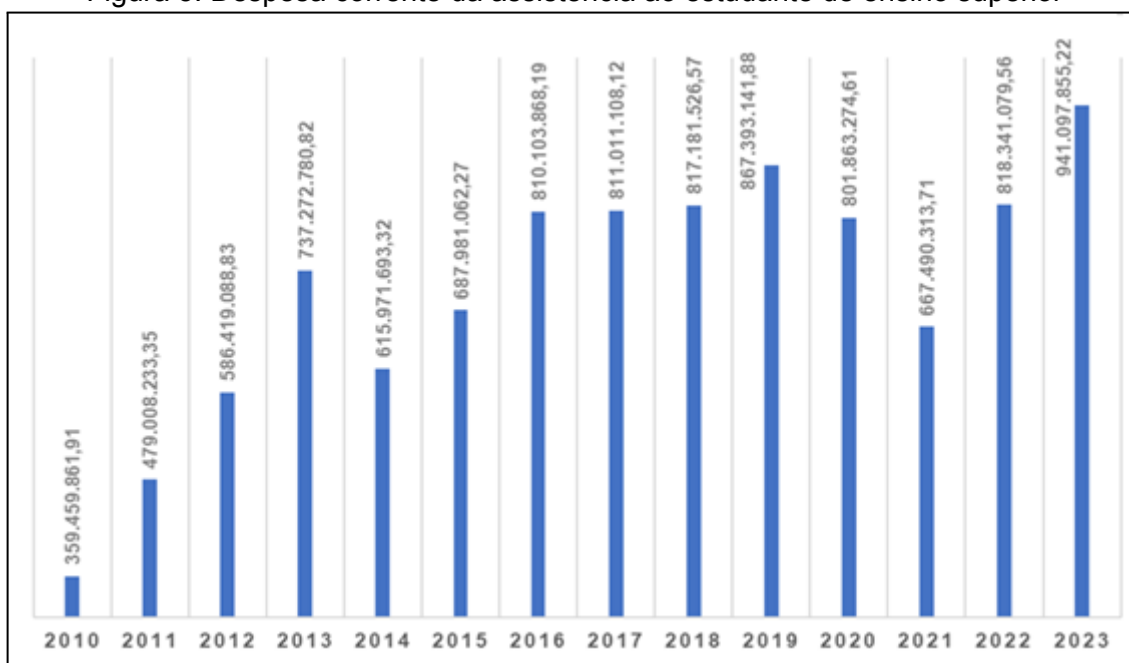
Em contrapartida, [Melo \(2019\)](#) adverte que, apesar de amplo em sua proposta de atuação, uma grave deficiência do PNAES é não adotar critérios mensuráveis que permitam monitorar o cumprimento dos objetivos indicados, bem como a falta de previsão regulamentar que habilite o controle social sobre a forma de utilização dos recursos

direcionados ao programa. Os estudos de [Dias e Carvalho \(2021\)](#) e [Lima e Ferreira \(2016\)](#) corroboram esta preocupação, afirmando que o processo de avaliação dos resultados da implementação do PNAES nas IFES acontece de forma isolada e superficial, não sendo possível, portanto, construir generalizações que habilitem a compreensão do alcance do programa em âmbito nacional. [Barbosa, Ferreira e Rezende \(2020\)](#) salientam que o Decreto do PNAES prevê que as próprias instituições definam quais mecanismos serão utilizados para acompanhamento e avaliação do programa no âmbito institucional, dificultando, assim, a realização de estudos comparativos entre os resultados alcançados por diferentes IFES.

Outro ponto em que o PNAES reforça a autonomia das instituições refere-se aos critérios para seleção dos estudantes contemplados com os auxílios, desde que sejam priorizados aqueles advindos da rede pública de educação ou que comprovem auferir renda *per capita* familiar limitada a um salário mínimo e meio ([BRASIL, 2010](#)). Esta concessão se estende, ainda, quanto à forma de administração e operacionalização dos recursos financeiros distribuídos pelo Governo Federal, permitindo sua aplicação conforme as particularidades de cada instituição e do perfil de seus estudantes ([SANTOS e FREITAS, 2014](#); [PEREIRA e ALMEIDA, 2015](#)).

Para [Lima \(2017\)](#), a importância do PNAES como instrumento facilitador da permanência estudantil e, conseqüentemente, do incremento no percentual da população qualificada por estudos em nível superior, deveria fazer com que o Governo Federal investisse maior volume de recursos não somente para distribuição de recursos financeiros diretos aos estudantes, mas também na contratação e qualificação de um corpo técnico de servidores capazes de atuar nos ambientes acadêmicos por meio de ações que promovam o atendimento àquelas necessidades estudantis não relacionadas com questões monetárias, como, por exemplo, saúde e interação social. Neste sentido, a Figura 5 apresenta a despesa corrente pública direcionada à assistência estudantil para estudantes de ensino superior entre 2010 e 2023. Nota-se grande crescimento nos recursos de 2010 para 2013, seguido de baixa em 2014 e retomada até 2019; em 2021 ocorreu uma drástica redução que levou o valor a um nível próximo ao observado em 2015, ainda que as IFES ofertassem maior quantidade de cursos e vagas e registrassem aumento nas matrículas de estudantes. Em 2023 houve a destinação do maior valor para as ações de assistência estudantil desde a publicação do Decreto de formalização do PNAES.

Figura 5: Despesa corrente da assistência ao estudante de ensino superior



Fonte: [Flores \(2022\)](#) e [Portal da Transparência \(2024\)](#).

Mesmo com a despesa anual do Governo Federal a título de assistência estudantil se mantendo próxima de 1 bilhão de reais em 2022 e 2023, o percentual de estudantes atendidos com os auxílios não passa de 20%, conforme aponta a pesquisa de [Bianchi e Adamczyk \(2021\)](#). Na Tabela 2 é possível verificar que a maior parte dos estudantes de ensino superior no Brasil não recebe nenhum tipo de benefício estudantil, porém nota-se discreta evolução no quantitativo de atendidos recentemente. Os autores enfatizam, contudo, que a avaliação da cobertura da política de assistência estudantil não deve ser balizada apenas por este indicador, pois outros fatores impactam na concessão dos benefícios, como, por exemplo, os critérios adotados por cada instituição para definir os candidatos em situação de vulnerabilidade.

Tabela 2: Distribuição de auxílio estudantil nas IFES

Ano	Recebe auxílio estudantil				Total
	Não		Sim		
	Quant.	%	Quant.	%	
2013	1.093.798	76,9%	328.715	23,1%	1.422.513
2014	1.248.679	83%	255.704	17%	1.504.383
2015	1.259.220	82,2%	272.135	17,8%	1.531.355
2016	1.295.768	81,8%	287.691	18,2%	1.583.459
2017	1.368.897	83,2%	276.323	16,8%	1.645.220

2018	1.369.010	81,8%	304.208	18,2%	1.673.218
2019	1.332.521	79,2%	349.634	20,8%	1.682.155

Fonte: [Bianchi e Adamczyk \(2021\)](#)

Dada a grande demanda por auxílios socioeconômicos pleiteados por estudantes que se encontram dentro do critério mínimo de elegibilidade estipulado pelo PNAES, as IFES enfrentam dificuldades para atender plenamente o público vulnerável, conforme destacam [Abreu e Ximenes \(2021\)](#), [Braga e Prá \(2021\)](#) e [Silva e Moreira \(2022\)](#). Desta forma, os gestores muitas vezes precisam adotar medidas que impactam diretamente nas condições que favorecem a permanência dos estudantes, como a redução do valor das bolsas, o cancelamento de contratos com funcionários admitidos exclusivamente para atendimento particular a estudantes com necessidades específicas, o repasse de custos à comunidade acadêmica (como nos Restaurantes Universitários, por exemplo) ou a supressão da oferta de serviços de saúde nos *campi* com estrutura para tal ([MARTINS, 2022](#); [SOARES, 2022](#)).

Ainda que não seja unanimidade entre os pesquisadores do tema, a institucionalização do PNAES pelo Governo Federal tornou possível a elaboração e consecução de ações que contribuíram (e continuam auxiliando, mesmo com as dificuldades orçamentárias atuais) para a manutenção de condições propícias à permanência de estudantes de cursos de graduação, bem como fomentaram a estruturação organizacional das IFES quanto aos setores de assistência estudantil. Vem sendo possível, portanto, oferecer no âmbito acadêmico um conjunto diferenciado de serviços para estudantes que realmente não dispõem de todos os meios, sejam materiais, pessoais ou financeiros, mas persistem na busca pela melhoria de sua qualificação como indivíduos e profissionais ([BIANCHI e ADAMCZYK, 2021](#); [SALDANHA e CASTRO, 2021](#)).

A seguir, apresenta-se um panorama sobre a estruturação da assistência estudantil no IFMG, abordando especialmente características relacionadas com a distribuição dos auxílios socioeconômicos aos estudantes da instituição.

2.2.1 Caracterização da assistência estudantil no IFMG

As informações utilizadas para o desenvolvimento desta Seção foram obtidas, em sua maioria, a partir do *website* institucional da Diretoria de Assuntos Estudantis (DIRAE) do IFMG ([IFMG, 2022b](#)) e na Política de Assistência Estudantil (PAE) da instituição ([IFMG, 2020](#)).

No que tange à estrutura organizacional, a DIRAE está inserida no âmbito da Pró-Reitoria de Ensino e funciona na Reitoria, tendo um servidor nomeado como Diretor

para atuar exclusivamente nos assuntos relacionados com a assistência estudantil da instituição. Essa Diretoria é auxiliada pelo setor de Coordenação de Assistência Estudantil (CAE) ou similar existente em cada Campus. Todas as ações da DIRAE e desses setores nos *campi* são pautadas pelos princípios e diretrizes contidos na PAE do IFMG, cuja última atualização ocorreu em 2020 com a publicação da Resolução Nº 9 de 03 de julho de 2020. Anteriormente denominada como Programa de Assistência Estudantil, a PAE observa incondicionalmente os preceitos advindos do PNAES, visando auxiliar a comunidade acadêmica no estabelecimento de ações e na elaboração de projetos que contribuam em diversos aspectos da vivência daqueles estudantes com demanda identificada. É válido mencionar que as ações da PAE estão disponíveis somente para estudantes regularmente matriculados nos cursos presenciais da instituição, com exceção daqueles em nível de Pós-Graduação, Formação Inicial Continuada (FIC) e extensão.

Os princípios e diretrizes que regem a PAE são amplos e abrangem desde propriedades institucionais até itens específicos vinculados aos estudantes. Quanto aos princípios, observa-se que as ações ancoradas na PAE devem, de maneira geral, defender a justiça social, respeitar a diversidade entre os indivíduos, promover a inclusão social por meio da educação, viabilizar a formação integral dos estudantes (considerando não somente o campo profissional, mas também o pessoal) e favorecer uma gestão participativa pautada pela transparência no uso dos recursos públicos. As diretrizes norteadoras da Política tratam, dentre outros aspectos, sobre a garantia de oferta de ensino gratuito e de qualidade, a definição de instrumentos aptos a incrementarem o controle social quanto às ações de assistência estudantil empreendidas pelo IFMG, a construção de espaços que integrem a escola com a família e a comunidade dos estudantes e o combate a quaisquer formas de discriminação e preconceito. Destaca-se, ainda, uma diretriz que reforça a necessidade de realizar pesquisas que subsidiem o planejamento das ações da assistência estudantil no âmbito da PAE, de maneira a promover a participação da comunidade acadêmica na definição das prioridades de atendimento, evitando-se também a execução de projetos sem a devida análise de aderência e viabilidade.

A PAE é baseada em três objetivos a serem alcançados com as ações da assistência estudantil no IFMG. O primeiro deles entende que esta política deve contribuir diretamente para a permanência dos estudantes matriculados nos cursos da instituição, ao mesmo tempo em que colabora com a redução das taxas de evasão e da desigualdade social existente na sociedade como um todo. No escopo do segundo objetivo, as ações da PAE precisam estar alinhadas, a fim de que seja fomentada a oferta de apoio pedagógico aos estudantes com dificuldades de aprendizado, momentâneas ou permanentes, favorecendo assim o incremento no desempenho acadêmico e a diminuição das taxas de retenção. O terceiro objetivo reforça alguns pontos abordados nas diretrizes da PAE ao indicar a

necessidade das ações de assistência estudantil serem concebidas e conduzidas a partir de um ambiente que preveja amplas condições de participação democrática.

Os programas da assistência estudantil do IFMG contam com recursos específicos alocados no orçamento institucional e são organizados em três categorias: programas de caráter universal, programas de apoio pedagógico e programas de caráter socioeconômico. Nota-se, portanto, a amplitude de abrangência da PAE, visando promover suporte financeiro e auxílio em diversas questões da vida acadêmica e pessoal dos estudantes.

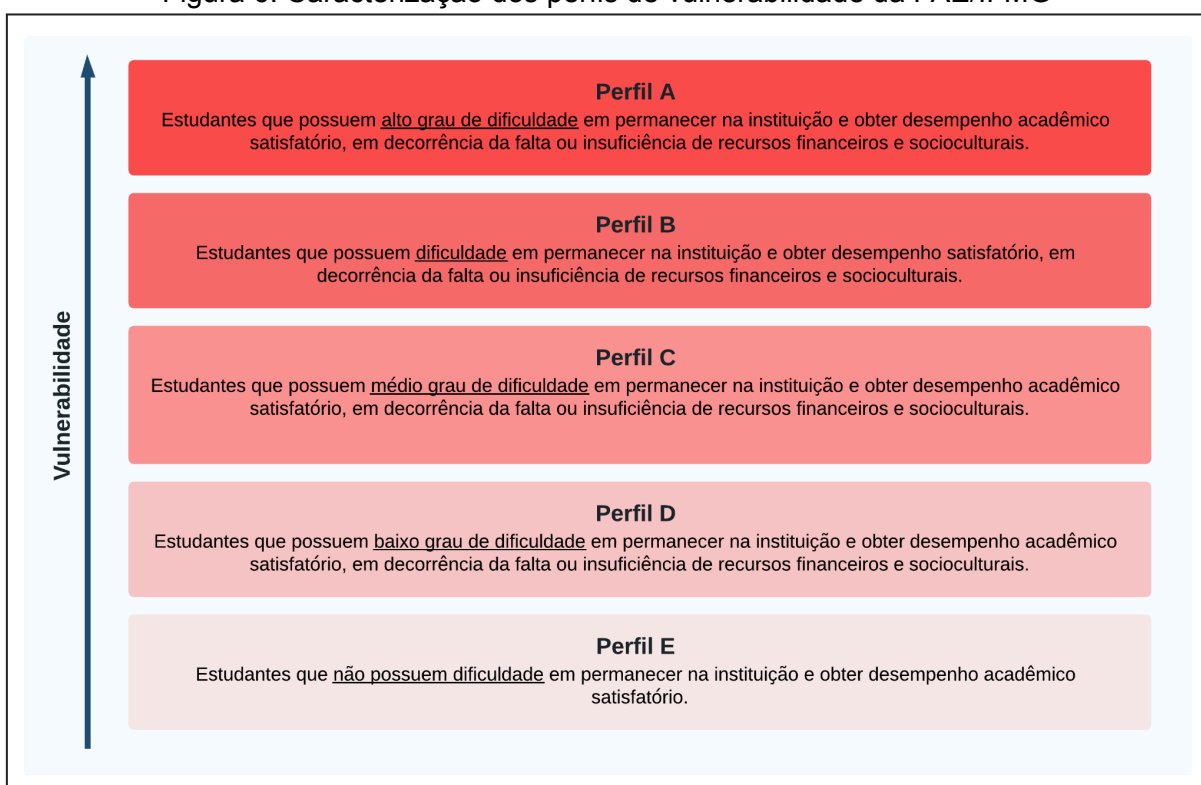
Os programas de caráter universal estão relacionados com as temáticas de acompanhamento e suporte ao ensino, assistência à saúde dos estudantes, incentivo à participação em ações acadêmicas/eventos e apoio a estudantes com necessidades educacionais específicas. Os programas de apoio pedagógico buscam atuar na complementação das demandas de formação acadêmica dos estudantes por meio da distribuição de bolsas de monitoria para disciplinas dos cursos técnicos e superiores. Os programas de caráter socioeconômico almejam contribuir com a solução de eventuais dificuldades financeiras dos estudantes e são implementados no IFMG através do repasse direto de recursos a estes, seja no formato de Bolsa Permanência (BP), isenção de pagamento no Restaurante Escolar ou quarto na Moradia Estudantil (nos *campi* que disponibilizam essa estrutura). Em virtude da pandemia de COVID-19, criou-se também um programa específico para auxílio emergencial no qual foram distribuídos recursos financeiros para que os estudantes pudessem adquirir equipamentos eletrônicos e participarem das aulas no formato remoto.

Como o foco desta tese reside na distribuição de auxílio socioeconômico e considerando o fato de que a maior parte dos *campi* do IFMG não disponibiliza restaurante e moradia para os estudantes, o programa de bolsa permanência será detalhado de maneira a destacar sua importância para o alcance dos objetivos da PAE. Trata-se de um programa global da instituição, isto é, oferecido em todos os *campi*, visando distribuir auxílio financeiro aos estudantes em vulnerabilidade que comprovarem essa situação dentro do contexto de processos seletivos regidos por edital próprio. Os valores das bolsas são escalonados em quatro níveis para que sejam atendidos diferentes perfis de vulnerabilidade socioeconômica: BP1 (R\$400,00), BP2 (R\$300,00), BP3 (R\$200,00) e BP4 (R\$150,00). O período de vigência das bolsas é de até dez meses durante o ano letivo.

A participação nos processos seletivos envolve uma análise conduzida pelo Núcleo de Assistentes Sociais do IFMG (NASIFMG), objetivando validar documentalmente as informações prestadas pelos estudantes que pleiteiam os auxílios no intuito de caracterizar aqueles com efetiva vulnerabilidade. A quantidade de bolsas distribuídas varia em função da disponibilidade orçamentária de cada Campus e do perfil dos candidatos. Eles são classificados em uma escala com cinco perfis de prioridade (“A” a “E”, sendo “A” o mais

prioritário e assim por diante) de acordo com o Índice de Vulnerabilidade Social (IVS), uma métrica calculada com base em dez indicadores obtidos a partir dos dados do candidato e de seu grupo familiar. São analisados dados da composição e renda familiar, bens patrimoniais, situação do trabalho e da moradia do estudante e do seu grupo familiar, procedência e meio de transporte utilizado para deslocamento ao Campus, além de questões de saúde ou necessidades específicas. A Figura 6 caracteriza os cinco perfis de vulnerabilidade definidos pela PAE do IFMG.

Figura 6: Caracterização dos perfis de vulnerabilidade da PAE/IFMG



Fonte: Elaborada a partir das informações de [IFMG \(2020\)](#)

No que tange ao acompanhamento dos estudantes atendidos por programas de caráter socioeconômico, as determinações da PAE são bem sucintas, mencionando que cada Campus deverá promover um monitoramento constante da frequência e do desempenho acadêmico dos assistidos e encaminhar aqueles com dificuldades de aprendizado para participação nos programas de apoio pedagógico. O auxílio poderá ser cancelado ou suspenso caso o estudante não participe de tais programas ou não cumpra as recomendações direcionadas pela Coordenação de Assistência Estudantil do Campus. Quanto à avaliação da PAE, está prevista análise anual envolvendo a DIRAE, a CAE dos *campi* e representantes dos estudantes, quando são observadas informações a respeito do alcance quantitativo dos programas em conjunto com as taxas de evasão/retenção, médias do desempenho acadêmico e frequência dos estudantes atendidos.

Os resultados anuais dos indicadores gerenciais do IFMG, incluídos aqueles dos programas de assistência estudantil, são publicados em documentos intitulados “Relatório de Gestão” ([IFMG, 2022c](#)). Além disso, dados históricos estão disponíveis na Plataforma Nilo Peçanha ([BRASIL, 2024](#)), permitindo traçar um panorama dos resultados quantitativos da assistência estudantil na instituição. A Tabela 3 foi elaborada com dados relativos ao período entre 2017 e 2023 e provê uma visão geral de vários indicadores que permitem observar, dentre outras ponderações, que o alcance do programa de bolsa permanência no IFMG ainda pode ser aprimorado.

Tabela 3: Distribuição de bolsa permanência e recursos empenhados (em R\$)

Indicador	2017	2018	2019	2020	2021	2022	2023
Matrículas	15.571	15.844	17.507	17.702	20.159	20.838	18.860
BP distribuídas	2.881	2.573	3.558	2.982	3.358	2.992	2.847
% atendidos	18,5%	16,2%	20,3%	16,8%	16,7%	14,4%	15,1%
Recursos empenhados para a AE	8.668.894,00	7.256.419,82	9.103.912,33	9.308.768,00	6.659.926,94	8.257.036,40	15.066.340,00
Varição anual dos recursos empenhados	0%	-16,3%	25,5%	2,3%	-28,5%	24,0%	82,5%
Recursos empenhados para BP	7.225.460,50	6.404.800,00	7.635.000,00	6.399.000,00	5.886.159,00	5.694.462,25	6.925.352,32
Proporção recursos BP x AE	83,3%	88,3%	83,9%	68,7%	88,4%	69,0%	46,0%

Fonte: Elaborada a partir de dados coletados em [BRASIL \(2024\)](#), [IFMG \(2022c\)](#) e [IFMG \(2023\)](#)

Analisando os dados da Tabela 3, percebe-se considerável evolução no número de matriculados em cursos cujos estudantes estão aptos a pleitear os auxílios gerenciados pela PAE, quais sejam, de ensino técnico ou graduação. Ademais, verifica-se incremento no número de bolsas permanência concedidas, porém sem mudança no percentual total de atendidos em função do crescimento conjunto do número de matriculados. Na verdade, o percentual de estudantes contemplados com bolsa permanência diminuiu do início até o final da série retratada na Tabela 3. O resultado do IFMG neste indicador está próximo da média nacional de 20,8%, em 2019, indicada por [Bianchi e Adamczyk \(2021\)](#).

Quanto aos recursos financeiros empenhados para os programas da assistência estudantil, nota-se variação não linear no período analisado, destacando-se a vultosa defasagem em 2021, justamente no momento em que a instituição possuía o maior número de estudantes matriculados da série e muitos deles se tornaram vulneráveis socioeconomicamente em razão da pandemia de COVID-19. Analisando o valor dos

recursos empenhados em contraste com o valor destinado ao programa de bolsa permanência, é possível perceber sua importância para a instituição, mesmo atendendo somente cerca de 20% dos estudantes matriculados. Em 2018 e 2021, por exemplo, os valores destinados a este programa corresponderam a 88% de todo o recurso financeiro empenhado para a assistência estudantil do IFMG. A variação deste indicador observada em 2020 se deve à realocação de recursos a fim de custear a oferta de auxílios temporários para aquisição de equipamentos tecnológicos, como notebooks e celulares, que permitissem aos estudantes participarem das atividades acadêmicas realizadas remotamente. Em 2022, a variação ocorreu em função do subsídio provido para a alimentação de todos os estudantes da Educação Básica matriculados na instituição.

2.3 Desempenho acadêmico

Nesta Seção é realizada a conceituação do construto desempenho acadêmico, permitindo sua compreensão dentro do escopo da pesquisa. São feitas, ainda, considerações sobre possíveis variáveis que atuam como determinantes do desempenho acadêmico, isto é, têm a capacidade de influenciá-lo positiva ou negativamente.

A forma como este construto é entendido, no nível teórico, apresenta relação direta com aspectos da sua mensuração e como tal processo é conduzido ([MAMEDE et al., 2015](#)). Neste sentido, [Munhoz \(2004\)](#) afirma que o desempenho acadêmico é o resultado de ações conduzidas por determinado estudante no ambiente institucional e avaliadas pelos educadores em termos de rendimento e eficiência, sendo comumente expresso no formato de conceitos ou notas. O desempenho seria, portanto, consequência de uma ou mais ações empreendidas pelo estudante durante o seu percurso acadêmico. Contudo, a autora ressalta que o nível do resultado medido como desempenho pode sofrer impactos de variadas situações, sejam elas particulares do estudante ou advindas da própria conjuntura institucional ou social.

Entendimento similar é apresentado por [Fineburg \(2009, apud Sousa, 2010\)](#) ao definir desempenho acadêmico como sinônimo de uma medida capaz de indicar o progresso do estudante no ambiente escolar, seja ela baseada em pontuações obtidas com a aplicação de testes específicos ou a partir do cálculo da média final de uma disciplina ou curso. A autora também salienta que numerosos fatores são capazes de interferir no processo vivenciado pelo estudante para alcançar o seu desempenho, tais como sua motivação pessoal e atribuições cotidianas.

Analisando a questão sob um ponto de vista educacional mais abrangente, [Fagundes, Luce e Espinar \(2014\)](#) afirmam que o desempenho acadêmico pode ser conceituado como um dos resultados da aprendizagem dos estudantes, advinda diretamente

do contexto vivenciado por eles em conjunto com seus professores e com a instituição de ensino. No entanto, os autores promovem uma ressalva quanto ao fato de que a aprendizagem é obtida a partir de uma variedade de formas e possibilidades, não apenas originada pela ação docente. Para [Silva et al. \(2015\)](#), o desempenho acadêmico engloba a avaliação sobre o aprendizado e o desenvolvimento alcançados pelos estudantes, permitindo analisar o cumprimento dos objetivos acadêmicos propostos quando do início do curso e, indiretamente, fornecendo subsídios que permitem compreender a efetividade da instituição como provedora de um ensino de qualidade.

Entretanto, autores como [Cordeiro e Silva \(2012\)](#) e [Nogueira \(2012\)](#) ponderam sobre os riscos de considerar o desempenho acadêmico apenas como o resultado de uma métrica sem contextualizá-la adequadamente aos estudantes. Por exemplo, uma das medidas mais usuais de indicação de desempenho são as notas finais obtidas nas disciplinas cursadas pelos estudantes. Logo, é fundamental que os processos de ensino e aprendizagem estejam alinhados com o formato de sua avaliação, a qual dará origem à medida de desempenho do estudante. Esta adequação contribui para evitar ruídos entre o entendimento dos envolvidos sobre as características necessárias das ações que, por sua vez, conduzirão a um desempenho satisfatório.

[Miranda et al. \(2015\)](#) ressaltam que existem muitas formas de mensurar o desempenho acadêmico, desde aquelas de operacionalização mais simples, como a nota ou conceito de determinada avaliação aplicada para os estudantes, até outras que envolvem maior complexidade para obtenção de uma medida final, como exames externos à IES ou o coeficiente de rendimento (CR). Para [Ferreira e Crisóstomo \(2012\)](#), a mensuração do desempenho acadêmico em um curso superior requer a definição de uma medida que considere a amplitude de tal nível de ensino, sendo que o CR consegue espelhar adequadamente esta realidade por permitir a integração de vários indicadores, como a carga horária matriculada e cursada pelo estudante, a quantidade de disciplinas cursadas, o seu aproveitamento em cada uma delas (seja este dado na forma de conceito ou numérica), o percentual de frequência individual e global, dentre outros. Ademais, trata-se de uma medida obtida ao final de cada semestre letivo, momento em que os indicadores são calculados, cumulativamente, em relação aos semestres anteriores, considerando, portanto, todo o histórico de desempenho do estudante ao longo do seu curso, e não somente no escopo de determinado recorte, como no caso de uma disciplina específica. De acordo com [Biembengut, Pacheco e Coninck \(2018\)](#), o CR é a medida usualmente adotada nas IFES para mensuração do desempenho acadêmico dos estudantes.

A forma de cálculo do CR pode variar em função da instituição ou, até mesmo, devido às particularidades de um curso de graduação. No Instituto Federal do Ceará (IFCE) e na Universidade Federal da Bahia (UFBA) o CR é calculado no encerramento de cada

período letivo e consiste de uma média das notas obtidas pelo estudante em cada disciplina cursada, ponderada pelos créditos que a disciplina equivale na matriz curricular ([JUCÁ et al., 2019](#); [RIBEIRO, PEIXOTO e BASTOS, 2017](#)). O cálculo do CR na Universidade Federal de Uberlândia (UFU) também é feito de forma cumulativa aos semestres letivos anteriores, encontrando uma proporção entre o produto das notas de cada disciplina e a carga horária cursada e a carga horária total matriculada ([QUEIROZ et al., 2015](#)). A Universidade Estadual Paulista (UNESP) adota como base de cálculo do CR apenas a média aritmética das notas finais de todas as disciplinas cursadas pelo estudante ao longo de sua trajetória acadêmica ([GALHARDO et al., 2020](#)). Na Universidade Tecnológica Federal do Paraná (UTFPR) o CR é apurado semestralmente com a aplicação de uma fórmula que encontra a média de todas as notas finais obtidas pelo estudante, ponderada pela carga horária total cursada até o momento ([BIEMBENGUT, PACHECO e CONINCK, 2018](#)).

No caso dos cursos de graduação do IFMG, o CR é obtido sob a forma de duas métricas: CR semestral e CR global. Enquanto o CR semestral consiste na média das notas finais, com ponderação pela carga horária daquelas disciplinas que o estudante cursou no semestre letivo, o CR global é a média das notas de todas as disciplinas cursadas pelo estudante até o momento do seu cálculo, ponderada pela carga horária cursada ([IFMG, 2018](#)).

2.3.1 Determinantes do desempenho acadêmico

Para [Nogueira \(2012\)](#), ao mesmo tempo em que se concebe o entendimento sobre a conceituação do desempenho acadêmico e suas múltiplas formas de mensuração, é necessário pontuar que trata-se de um construto que pode sofrer influência de fatores que impactam no processo de aprendizagem e, conseqüentemente, nos resultados do ensino. [Brown et al. \(2008\)](#) corroboram este entendimento ao afirmarem que o desempenho dos estudantes, muitas vezes, é determinado ou afetado por um complexo conjunto de variáveis que precisam ser consideradas paralelamente à análise do indicador medido a título de desempenho acadêmico, tais como características pessoais e familiares dos envolvidos, organização estrutural da IES, aspectos dos docentes e especificidades do curso superior.

Essas variáveis que contribuem, negativa ou positivamente, para o desempenho acadêmico, são categorizadas por [Guney \(2009\)](#), [Ferreira \(2015\)](#) e [Miranda et al. \(2015\)](#) em três dimensões, relacionadas ao corpo discente, ao corpo docente e à IES. Outros autores, como [Gomes e Soares \(2013\)](#) e [Sousa, Ferreira e Miranda \(2019\)](#), agrupam essas variáveis determinantes dadas suas relações com aspectos sociais, econômicos e culturais do estudante e da instituição. Sob qualquer um desses aspectos, conforme apontam [Almeida e Lopes \(2020\)](#) e [Teles et al. \(2019\)](#), é mais usual que um conjunto de variáveis ocasione oscilações no desempenho acadêmico, salvo em casos específicos nos quais somente uma

variável impacta os resultados obtidos por determinado estudante. Por consequência, a visão das variáveis integradas em algum agrupamento se mostra mais adequada para compreender os fatores que determinam o desempenho em determinada instituição.

Diferentes estudos científicos encontraram indícios que apontam algumas variáveis relacionadas (positiva ou negativamente) ao desempenho acadêmico. A Tabela 4 organiza algumas dessas variáveis de acordo com a sua categorização em uma das três dimensões mencionadas anteriormente: corpo discente, corpo docente e IES.

Tabela 4: Variáveis relacionadas com o desempenho acadêmico

Dimensão	Variáveis
Corpo discente	Gênero, idade, etnia, estado civil, escolaridade dos pais, nível de renda da família, desempenho escolar anterior, estar trabalhando, horas de estudo, horas de sono, participar de projetos institucionais.
Corpo docente	Titulação, experiência profissional, regime de trabalho, estratégia ou método de ensino.
IES	Disponibilidade de recursos tecnológicos, infraestrutura dos ambientes de estudo, recursos didáticos, materiais disponibilizados aos alunos, tamanho das turmas.

Fonte: Adaptado de [Ferreira e Crisóstomo \(2012\)](#), [Miranda et al. \(2015\)](#), [Paiva e Lourenço \(2011\)](#) e [Rodrigues et al. \(2016\)](#)

Quanto às variáveis da dimensão do corpo discente, [Moreira e Gomes \(2018\)](#) encontraram relação positiva entre o gênero feminino e o desempenho acadêmico, assim como [Almeida, Cassuce e Cirino \(2016\)](#) e [Rangel e Miranda \(2016\)](#). Esses autores observaram, ainda, que a quantidade de reprovações do estudante impacta no seu desempenho ao longo do curso, fator que indica a necessidade de monitoramento pelos gestores acadêmicos no sentido de promover ações institucionais capazes de auxiliar aqueles com dificuldades na condução da sua vida escolar. [Santos, Figueiredo e Vieira \(2018\)](#) identificaram relação positiva entre a assiduidade dos estudantes nas atividades acadêmicas e o desempenho obtido ao final do semestre letivo. Quando analisada a idade dos estudantes, foi encontrada por [Casiraghi et al. \(2020\)](#) e [Rodrigues et al. \(2016\)](#) variação significativa do desempenho acadêmico, com destaque para o maior aproveitamento obtido por aqueles mais jovens.

As relações entre a renda familiar (incluída a do próprio estudante) e o desempenho acadêmico foram analisadas por [Araújo, Ferreira e Silva \(2022\)](#), [Brandt \(2017\)](#), [Ferreira et al. \(2021\)](#) e [Saraiva et al. \(2021\)](#), indicando que esta variável é capaz, nos contextos analisados, de impactar o resultado final dos discentes na medida em que a renda auferida é maior. Tal relação pode ser explicada pelas diversas possibilidades trazidas por núcleos familiares que dispõem de mais recursos para sua sobrevivência. Nestes casos, o estudante

não precisa trabalhar a fim de se manter no curso, possui mais tempo disponível para dedicação aos estudos e para a realização de atividades acadêmicas, pode se deslocar com mais facilidade ao Campus e adquirir materiais adequados que incrementam seus conhecimentos, dentre outras condições oportunizadas por um maior nível de renda familiar.

Os resultados acadêmicos anteriores ao ingresso no curso superior também podem ser associados ao desempenho dos estudantes, segundo [Lima e Pazinato \(2018\)](#) e [Sousa, Ferreira e Miranda \(2019\)](#). Esses autores identificaram, por exemplo, relação direta entre a formação geral obtida no Ensino Médio e o rendimento final na graduação, isto é, quanto menores as notas dos egressos em seus cursos de origem, menores são os coeficientes de rendimento no ensino superior. A qualidade e a quantidade das horas de sono do estudante e sua relação com o desempenho foram analisadas por [Queiroz \(2021\)](#), tendo sido observada significância positiva entre aqueles indivíduos que possuem boa qualidade de sono. [Niquini et al. \(2015\)](#) analisaram as relações entre a carga de trabalho do estudante e o seu rendimento no curso superior, observando que maiores jornadas de trabalho podem contribuir para a piora no desempenho acadêmico. [Araújo, Andriola e Coelho \(2018\)](#) e [Nardini et al. \(2019\)](#) encontraram evidências de que a participação em projetos institucionais, em especial aqueles de iniciação científica, afetam positivamente o desempenho acadêmico de estudantes de cursos de graduação.

Alguns fatores advindos das características do corpo docente da instituição também foram analisados por estudos que identificaram sua relação com o desempenho acadêmico. [Martins e Marinho \(2019\)](#) e [Miranda et al. \(2015\)](#) encontraram correlação estatisticamente significativa entre a estratégia ou método de ensino adotado pelo docente e o rendimento dos estudantes, sobretudo quando técnicas expositivas são adotadas como abordagem principal para a condução das aulas. A titulação acadêmica dos docentes também foi identificada como fator positivo para o desempenho geral dos cursos (medido através dos resultados dos estudantes em avaliações específicas), de acordo com [Ferreira \(2015\)](#) e [Miranda, Nova e Cornacchione Júnior \(2013\)](#). O regime de trabalho dos docentes, em especial aqueles vinculados à dedicação exclusiva, também auxilia na explicação do desempenho de estudantes de cursos superiores, conforme apontado por [Santos \(2012\)](#).

Sobre as variáveis relacionadas com a IES, [Anegues, Porto Júnior e Figueiredo \(2020\)](#) encontraram relação direta e negativa entre a quantidade de estudantes na turma e o seu desempenho, afirmando que o aumento de matriculados tende a reduzir o rendimento acadêmico na média geral, especialmente nas turmas de ingressantes. [Pavão, Ferreira e Ferreira \(2019\)](#) apresentam um estudo que identificou dois fatores internos das instituições capazes de influenciar positivamente o desempenho dos estudantes: a atenção recebida por eles durante as suas diversas interações com a IES ao longo do seu percurso acadêmico, seja com os funcionários administrativos, coordenação de curso ou docentes, e o projeto

pedagógico do curso em questão, englobando a organização e integração das disciplinas, a carga horária dos componentes curriculares e sua aderência à realidade dos estudantes. [Ferreira \(2015\)](#) e [Martins e Marinho \(2019\)](#) destacam a importância dos ambientes de estudos disponibilizados pela IES para o corpo discente ao apresentarem resultados demonstrando que os estudantes que utilizam tais espaços pelo menos duas vezes por semana possuem melhor desempenho.

É válido ressaltar que o impacto no desempenho acadêmico, influenciado pelas variáveis citadas, não necessariamente ocorreu em todos os estudos analisados, isto é, em cada contexto o desempenho foi afetado de maneira diferente pelas variáveis, ou sequer foi afetado, dado o perfil multivariado dos estudantes envolvidos. Esta constatação reforça a importância da realização de estudos que, focados na realidade de determinada IES, consigam indicar com clareza e validade estatística quais são os fatores capazes de contribuir positivamente para o incremento do desempenho acadêmico dos estudantes de cursos de ensino superior. A tomada de decisões gerenciais baseadas em estudos conduzidos com cenários diferentes daquele da IES pode se mostrar ineficaz do ponto de vista dos resultados esperados com as ações planejadas e implementadas, pois nem sempre a generalização dos achados em uma avaliação científica é adequada para cenários que envolvem aspectos multifatoriais, como é o caso do desempenho acadêmico ([CALIATTO e ALMEIDA, 2020](#); [CASIRAGHI et al, 2020](#)).

2.4 Gestão da Informação (GI)

O propósito desta Seção é fornecer uma breve contextualização da área de Gestão da Informação, abordando a conceituação dos seus principais temas, analisando o papel da Tecnologia da Informação nos processos modernos de GI e compreendendo a importância do conhecimento sobre o comportamento informacional dos indivíduos vinculados a uma organização.

De acordo com [Wilson \(2002\)](#), uma conceituação ampla da Gestão da Informação a compreende como um processo contínuo e sistematizado que visa obter, organizar, gerar, distribuir e empregar informações que atendam as demandas organizacionais identificadas preliminarmente. [Choo \(1995\)](#) segue linha de raciocínio similar, indicando que as atividades que preveem a criação, obtenção, armazenamento, compreensão e utilização das informações no contexto organizacional constituem a base intelectual que permite a evolução das organizações inteligentes por meio do aprendizado constante e sistematizado. [Detlor \(2010\)](#) também conceitua a Gestão da Informação elencando as mesmas atividades sobre as informações citadas por Wilson e Choo, adicionando que o objetivo final de tal gerenciamento é que as organizações tenham processos bem definidos para acesso,

processamento e uso da informação em seus ambientes. Uma análise histórica dos aspectos evolutivos da Gestão da Informação como área de estudos foi conduzida com abrangência por [Martins \(2014\)](#), onde é possível não somente identificar a definição de diferentes conceituações sobre o tema, mas principalmente compreender sobre a mudança na percepção do valor da informação para as organizações modernas.

[Oliveira et al. \(2014\)](#) compreendem a GI como um agrupamento de processos que atuam em consonância para suportar as atividades de aprendizagem de uma organização. Neste aspecto, a informação se mostra como uma matéria-prima fundamental para que o aprendizado sirva aos propósitos de tomada de decisão, demandando, portanto, estruturas e ferramentas que garantam a sua qualidade. Entendimento similar é encontrado em Davenport (2004) quando é afirmado que a quantidade de informações disponíveis para a tomada de decisão já não se mostra mais como um problema, mas sim a dificuldade para que as informações utilizadas sejam originadas de um eficiente processo de gerenciamento que garanta sua correteza, tempestividade e integridade.

Certa concordância é observada nas afirmações de [Bergeron \(1996\)](#) e Dias e Belluzzo (2003) quando indicam que o foco da GI está diretamente ligado ao processo decisório, isto é, trata-se de uma área em que os conceitos, princípios, métodos e técnicas estão voltados para promover efetividade na tomada de decisão organizacional. [Valentim \(2002\)](#) entende a GI como um conjunto de estratégias que apoiam as decisões organizacionais mediante um amplo trabalho de entendimento das necessidades informacionais, sua coleta, armazenamento, tratamento e disponibilização para análises sob variados pontos de vista, cujos resultados suportarão a construção de novos conhecimentos dentro do ambiente da organização.

Tarapanoff (2001) adiciona que, para gerar resultados factíveis neste processo, a GI necessita, por meio de variadas ferramentas e meios, identificar adequadamente os recursos informacionais da organização, os quais poderão ser potencializados no sentido de fazer com que esta desenvolva novas capacidades de aprendizado e adaptação às mudanças. Uma ressalva é feita por Davenport e Prusak (1998) ao ponderarem que o alcance de uma decisão efetiva baseada em estratégias informacionais depende, em grande parte, do atendimento às características básicas da informação utilizada, quais sejam: integridade, atualização, precisão, valor e interpretabilidade.

Nonaka e Takeuchi (2008) destacam o fato de que a GI contribui estrategicamente para que ocorra, no âmbito organizacional, a conversão do conhecimento tácito, aquele proveniente da personalidade dos indivíduos, em conhecimento explícito, o qual pode ser compartilhado sob variados formatos, tais como dados, recursos visuais, manuais ou especificações. A formalização dos resultados deste processo de conversão auxilia, portanto, na garantia da aprendizagem contínua pela organização, que passa a ter

capacidade de difundir sistematicamente para outros indivíduos conhecimentos variados que, até então, estavam de posse de um número restrito de indivíduos. Sveiby (1998) ressalta que organizações com capacidade de replicação sistemática de conhecimentos conseguem apresentar, com maior frequência, diferenciais competitivos que colaboram para incrementar o seu valor de mercado.

Segundo [Barbosa \(2008\)](#), é preciso salientar que o fenômeno central investigado pela GI é o conhecimento explícito, diferente, por conseguinte, daquele trabalhado na Gestão do Conhecimento (GC), o conhecimento tácito. Enquanto a GI trata os diversos documentos produzidos e armazenados no âmbito organizacional, estando ligada estreitamente com os Sistemas de Informação, a GC lida com o conhecimento pessoal dos indivíduos, o qual muitas vezes sequer é formalizado e compartilhado. [Valentim \(2002\)](#) adita relatando que ambas as áreas atuam na elaboração de modelos de gestão com características complementares, não dissociativas, ficando a GI responsável por atuar mais diretamente com os fluxos formais da informação e a GC com os fluxos informais, isto é, abordando o que não está explicitado nos modelos.

Na visão de [Bergeron \(1996\)](#), os estudos relacionados com a GI podem ser analisados em função de duas perspectivas distintas, denominadas pela autora como tecnológica e integrativa. Na vertente tecnológica, ocorrem com frequência entendimentos equivocados de que informação é sinônimo de tecnologia e que a informação somente pode ser provida por sistemas informatizados. Por outro lado, a vertente integrativa prevê o entendimento da GI com uma perspectiva conexa sobre os problemas informacionais da organização, ou seja, trata-se de uma visão na qual a GI não envolve somente a tecnologia em si, mas busca ser uma referência quanto à integração harmônica entre as informações, os sistemas e serviços necessários para que os indivíduos estejam habilitados a acessarem recursos que contribuam para o alcance dos objetivos organizacionais.

No que tange ao maior aprofundamento da compreensão da GI, abarcando suas complexidades e especificidades, [Barbosa \(2016\)](#) e [Martins \(2014\)](#) entendem como insuficiente somente o estudo do histórico da área e a análise de seus conceitos. É preciso que a GI seja compreendida no âmbito de algum modelo de implementação composto de teorias e práticas conjuntas, além do tratamento do ciclo de vida da informação. Assim, torna-se factível conceber uma estrutura informacional dotada de elementos essenciais que promoverão resultados tangíveis para as organizações no sentido de fazer com que a informação tenha todas as características requeridas para ser considerada de qualidade.

[Pereira \(2011\)](#) alerta que o entendimento do fluxo informacional, necessário para a implementação eficaz da GI, normalmente se mostra como uma tarefa complexa dada a existência de vários canais transmissores (fontes) de informações no contexto organizacional. Ademais, essas fontes não permeiam este ambiente de maneira isolada,

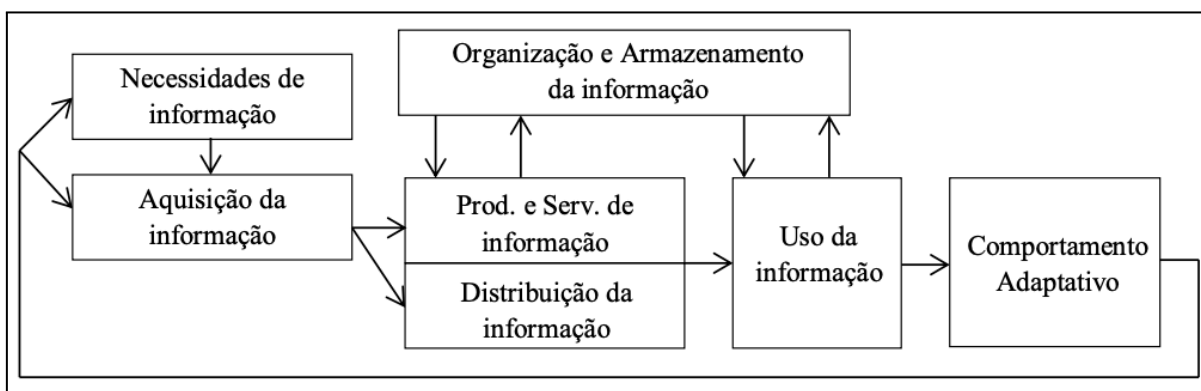
mas interagindo para processar e adicionar valor à informação final acessada pelos indivíduos, tendo capacidade, por conseguinte, de influenciar o seu comportamento informacional.

Dentre os modelos de implementação da GI concebidos ao longo dos últimos anos sob a vertente integrativa, destacam-se na literatura da área aqueles originados de McGee e Prusak (1994), Davenport (2002), Choo (2003) e Marchand, Kettinger e Rollins (2001). Há de se ressaltar que nenhum modelo pressupõe a ausência de adaptações prévias à sua implementação no ambiente organizacional. Ademais, dado o fato de que cada organização é única em suas características, a mesclagem de modelos pode contribuir para que os processos de GI estejam alinhados com as demandas organizacionais e dos indivíduos.

Esta tese não objetiva implementar um modelo de GI na organização estudada, mas em virtude da importância deste recurso para o entendimento dos conceitos mencionados nesta Seção, o modelo de Choo será brevemente apresentado. A aderência deste modelo à presente pesquisa reside no fato de que o mesmo avalia o gerenciamento da informação sob a ótica do resultado estratégico, isto é, não se trata apenas de um modelo genérico que normaliza o fluxo informacional da organização, mas enfatiza a retroalimentação como fundamento para o aprendizado, base para a geração de conhecimentos e inovação.

Em linhas gerais, o modelo de Choo se baseia na crença do autor de que a organização e seus indivíduos precisam aprender a todo momento para que, futuramente, o conhecimento gerado possa ser oportunamente empregado na resolução de algum problema ou no desenvolvimento de bens e serviços. O propósito geral da utilização do modelo é servir para que as informações identificadas como pertinentes ao negócio da organização estejam acessíveis quando demandadas. Outro ponto de destaque é que o valor à informação é atribuído pelos indivíduos, isto é, ela precisa demonstrar sentido para quem a recebe e ter significado atribuído, cumprindo assim as condições para ser, de fato, uma informação útil ao contexto organizacional. O ciclo informacional proposto por Choo é constituído por processos (exibidos na Figura 7) responsáveis para que o gerenciamento da informação seja de fácil compreensão pelos indivíduos e as informações transitadas pelo ambiente organizacional se transformem em percepção, conhecimento e ações.

Figura 7: Ciclo de gerenciamento da informação de Choo



Fonte: Adaptado de [Choo \(2002\)](#)

O primeiro processo visa identificar as necessidades informacionais da organização, de forma que elas transmitam representativamente a problemática enfrentada pelos usuários. Dado o fato de que tais necessidades estão, comumente, relacionadas a problemas ou tarefas conduzidas pelos usuários, dois aspectos podem colaborar para sua identificação: I) Qual é o objeto da necessidade, ou seja, que informação é demandada; II) Quais são as exigências do contexto, indicando as razões da necessidade e as formas de utilização da informação. Para que os resultados deste processo sejam eficazes à continuidade do ciclo, é imperativo que as necessidades levantadas estejam precisamente descritas e formalizadas.

O segundo processo trata da aquisição da informação, podendo se mostrar como um dos mais complexos de todo o ciclo, dada a ampla variedade de fontes de dados disponíveis atualmente. Choo questiona o fato de que, geralmente, as pessoas envolvidas com a organização não são consideradas como as fontes prioritárias de informações, posto ocupado por bancos de dados eletrônicos ou documentos impressos, a despeito de estarem em contato frequente com clientes e fornecedores, atuarem em variados tipos de projetos, bem como reunirem, ao longo do tempo, conhecimentos especializados e experiência sobre o trabalho em si.

As atividades vinculadas com a organização e o armazenamento da informação são contempladas no terceiro processo, o qual também se mostra muito importante para o gerenciamento informacional e, ao mesmo tempo, com alto grau de complexidade para implementação e manutenção. Choo ressalta que é necessário equilibrar todas as demandas de informação com um adequado nível de flexibilidade, construindo um ambiente orientado não somente para o acesso aos documentos, mas também aos usuários. Outro ponto que requer atenção refere-se à qualidade da indexação dos documentos, requisito fundamental para que o processo de recuperação das informações seja eficiente.

Choo define um processo específico para o desenvolvimento de produtos e serviços da informação, presumindo que a mesma já se encontra devidamente organizada e armazenada. O objetivo é construir programas, práticas e procedimentos capazes de empacotar a informação de acordo com a demanda de cada interessado, customizando-a em diversos formatos como, por exemplo, relatórios, gráficos, esquemas e *hotsites*, conforme o perfil do público-alvo a ser atendido.

O processo de distribuição da informação é entendido por Choo sob duas perspectivas: compartilhamento e disseminação. Enquanto as atividades de compartilhamento auxiliam a organização no provimento de uma aprendizagem heterogênea e contextual, onde os *insights* individuais encontram formas de se transformar em inovação, as atividades de disseminação buscam fornecer conteúdos específicos para aqueles indivíduos que requerem informações qualificadas em seu cotidiano. O autor ressalta a importância deste processo não somente para o ciclo da GI, mas também como meio de ligação com as atividades de gestão do conhecimento que, porventura, estejam sendo conduzidas na organização.

O fluxo do ciclo informacional proposto por Choo culmina em um processo dedicado às atividades interligadas com o uso da informação, cuja ocorrência pode se dar tanto em função da necessidade de alguma tomada de decisão quanto para auxiliar no aprimoramento dos conhecimentos abrangidos no contexto de eventual situação problema. É fundamental que os indivíduos partam do pressuposto de que o uso da informação, dentro do âmbito do aprendizado organizacional, deve acontecer com o intuito de criar conhecimento, em especial aquele moldado como representações significativas e contextualizadas que ampare a execução de ações intencionais.

Por fim, é evidenciado no modelo que o ciclo informacional precisa de retroalimentação para que a capacidade de adaptação da organização seja desenvolvida e as condições ambientais para aprendizado estejam presentes. As necessidades de informação requerem reavaliações para que seja possível identificar se o modelo atual atende aos anseios dos indivíduos ou se o conhecimento corrente não se mostra suficiente para a tomada de decisões, indicando, portanto, que novos aprendizados devem ser construídos com o suporte informacional.

2.4.1 Relações entre a Tecnologia da Informação e a GI

Embora originada com preocupações documentais e de controle de fluxos procedimentais, a GI se desenvolveu fortemente nas últimas décadas compondo uma integração com recursos de Tecnologia da Informação que, por sua vez, contribuíram para que as políticas e procedimentos concebidos pelas organizações para gerir o ciclo informacional fossem efetivamente implementados ([CIANCONI, 2003](#)). Evidencia-se,

portanto, que as tecnologias não devem ser utilizadas apenas como um fim em si mesmas (DAVENPORT, 2002), mas sobretudo como auxiliares para os diversos processos organizacionais que permeiam a gestão estratégica como base para o sucesso empresarial (MCGEE e PRUSAK, 1994). [Malin \(2003\)](#) complementa destacando o papel do uso intensivo de Tecnologias de Informação no ambiente corporativo como um dos fatores que auxiliaram na consolidação da informação como um elemento demandante de gerenciamento formal.

[Savic \(1992\)](#) observa que a exploração dos novos recursos ofertados como tecnologias computacionais permitiu a reestruturação de diversas atividades no âmbito organizacional como, por exemplo, aquelas relacionadas com o armazenamento, recuperação e processamento de informações. Tais tecnologias, ao introduzirem inovadores formatos de manipulação de recursos informacionais e otimizarem inúmeras tarefas executadas manualmente até então, fizeram com que a rotina dos indivíduos fosse aperfeiçoada e, em algumas áreas, completamente alteradas, dando lugar a um ambiente no qual a tecnologia desempenha um importante papel para que a informação seja valorizada como ativo tanto para as organizações quanto para o público em geral.

Para [Macedo e Barbosa \(2013\)](#), o emprego massivo de tecnologias computacionais no escopo da GI de empresas dos mais variados ramos de negócios está ligado, em especial, com o surgimento cada vez mais frequente de ambientes concorrenciais que requerem dos gestores extrema capacidade decisória mesclando qualidade e agilidade. Na medida em que o alcance dos resultados de uma organização depende das decisões estratégicas e percepções de seus gestores, o emprego adequadamente planejado da TI consegue promover o uso eficaz da informação como subsídio para os processos decisórios.

Partindo da premissa de que os recursos de TI são implementados e utilizados como apoio à tomada de decisão organizacional, Marchand, Kettinger e Rollins (2000) entendem que as práticas de TI aplicáveis a tal processo podem ser categorizadas em quatro níveis. No nível mais alto, indicado como **apoio gerencial**, encontram-se os recursos de TI utilizados por indivíduos responsáveis pelas decisões estratégicas empresariais. No nível de **apoio à inovação**, os recursos de TI visam contribuir com o ferramental necessário para a exploração de novas ideias que promovam o desenvolvimento de novos serviços e produtos. O nível de **apoio à processos** abarca aquelas tecnologias que são empregadas para o planejamento de recursos. O último nível contempla o **apoio operacional** oferecido pela TI, particularmente o suportado por sistemas de informação gerenciais que viabilizam a informatização das rotinas operacionais da organização. É ressaltado pelos autores que, independentemente desta categorização, é fundamental que a organização construa suas práticas de TI de maneira que o seu emprego esteja integrado aos comportamentos e valores informacionais dos indivíduos que nela atuam ou se relacionam.

Ao buscarem construir uma definição mais atualizada para a GI enquanto área de pesquisa, [Nonato e Aganette \(2022\)](#) afirmam que ela não pode estar dissociada da TI, indicando que se trata de um dos três pilares nos quais a GI se baseia, conjuntamente com o ciclo de vida da informação e a gestão organizacional. Para os autores, assim como defendido por Castells (2011), a importância da TI para a implementação de políticas efetivas de GI no atual contexto organizacional é imensurável, de forma que a penetrabilidade cada vez mais ampla das tecnologias computacionais nos processos de trabalho que envolvem controle da informação faz com que este cenário seja irreversível.

[Barbosa \(2020\)](#) ressalta que a TI possui relevante parcela de contribuição no incremento da produção e transmissão da informação nas últimas décadas, permitindo que a disponibilidade deste recurso transformasse sob variadas óticas a gestão empresarial. Por outro lado, o autor adverte que a redução do hiato entre a disponibilização da informação e a capacidade de uso realista da mesma ainda é um desafio enfrentado pelos indivíduos nas organizações. É preciso que os processos organizacionais elaborados no âmbito da GI estejam em consonância com as necessidades dos indivíduos, reduzindo assim a percepção de sobrecarga informacional que, em muitos contextos, afeta as possibilidades de uso harmônico da informação.

Os recursos de TI são apresentados por [Pinto, Molina e Paletta \(2022\)](#) como os principais motivadores para o aumento da vantagem competitiva das organizações modernas cuja oferta de produtos e serviços é baseada em informação. A redução no tempo para processamento de atividades de informação, a oferta de formatos variados para armazenamento da informação, a ampliação da capacidade de recuperação da informação e a eficiência em compartilhá-la com um público a cada dia mais ávido por informações factíveis fazem da TI um parceiro ideal para colocar em prática as propostas da área de GI. A partir do momento em que a TI consegue suportar os processos de GI e a integração destes com atividades que visam a geração de conhecimentos, ocorre uma busca constante por parte dos administradores para encontrar tecnologias que tragam uma diferenciação para o tratamento da informação no escopo organizacional.

Um fator pouco explorado na literatura da área de GI é abordado por [Lopes \(2020\)](#) e [Gomes et al. \(2019\)](#) ao realçarem o vínculo de dependência das organizações com empresas fornecedoras de produtos e serviços de TI. A adoção de práticas da GI baseadas no uso intensivo de tecnologias computacionais que precisam estar disponíveis no esquema 24 x 7 (vinte e quatro horas por dia, sete dias da semana) faz com que os contratos assinados entre as partes fiquem, muitas vezes, desequilibrados a ponto de privilegiar os fornecedores, em especial no quesito valor. Trata-se de uma situação que deve ser adequadamente estudada e planejada sob o ponto de vista estratégico da organização,

possibilitando que a TI seja incorporada à estrutura funcional como um ativo e alinhada com as demandas institucionais.

[Araújo \(2016\)](#) salienta que, embora seja possível afirmar que a TI contribui para o incremento dos resultados de uma organização, particularmente quando integrada a um ambiente estruturado com o apoio da GI, existem fatores que podem limitar a sua aderência às estratégias do negócio em si. Dentre tais obstáculos destacam-se a resistência à mudanças, dificuldades com integração entre vários recursos tecnológicos, suporte técnico comprometido em termos de qualidade de atendimento, necessidades constantes de atualização dos recursos e empecilhos com fornecedores. Para o autor, um planejamento estruturado que englobe como a TI estará alinhada com as estratégias organizacionais tende a reduzir consideravelmente os riscos da sua adoção, bem como contribui para que o ciclo informacional consiga prover os recursos necessários aos gestores para auxiliar nos processos de tomada de decisão.

2.4.2 Comportamento informacional

De acordo com [Araújo e Paula \(2017\)](#), uma das preocupações mais marcantes da Ciência da Informação ao longo de sua trajetória enquanto área de conhecimento vincula-se com a busca pelo entendimento das características comportamentais das pessoas no que tange à necessidade, busca e uso da informação, fato que contribuiu para a elaboração de diversos modelos com este intuito.

Ainda sob esse aspecto evolucionar da área, [Gasque e Costa \(2010\)](#) mencionam que ele ocorreu não somente no que tange aos assuntos pesquisados, mas envolveu a mudança na orientação dos objetos de estudo, saindo de uma centrada em sistemas para uma direcionada aos usuários e, posteriormente, aos comportamentos destes quanto à informação. Isso fez com que tarefas ou atividades específicas de informação perdessem o foco principal para a compreensão da situação pessoal, organizacional ou social onde sucede o surgimento da informação e o local de sua utilização. Esses autores promovem em seu trabalho uma ampla caracterização histórica de como o conceito de comportamento informacional foi sendo construído ao longo do tempo.

Algumas definições mais genéricas sobre o termo comportamento informacional foram elaboradas por autores com o intuito de explorar toda a abrangência englobada em seu escopo. Davenport e Prusak (1998) afirmam que estudos nesta área buscam compreender como os indivíduos se relacionam com a informação, sendo que para isso é preciso analisar como acontecem os processos de busca, utilização, mudanças, acúmulo e eliminação da informação. Tarapanoff (2011) complementa que o comportamento informacional objetiva desvendar as formas de interação entre as pessoas e as informações que elas coletam por conta própria (atitude ativa) ou recebem a partir de terceiros (atitude

passiva), permitindo, assim, entender como essas interações afetam os diversos contextos da vida de tais pessoas. Para Todd (2003), o comportamento informacional retrata todo tipo de conduta humana que visa buscar e utilizar informações, cuja aplicação prática pode se dar na interação entre os indivíduos, na análise de dados brutos obtidos sob variados formatos e de diversas fontes ou no emprego da informação obtida para geração de novas ideias e conhecimentos, fazendo com que as ações e decisões sejam precedidas de um processo composto por atividades que, em última instância, contribuem para ampliar os conhecimentos prévios dos indivíduos e organizações.

A importância da noção de contexto como elemento fundamental para a compreensão dos comportamentos informacionais dos indivíduos foi reforçada por [Pettigrew, Fidel e Bruce \(2001\)](#). Para os autores, trata-se de uma área de estudos que requer a percepção das necessidades dos indivíduos e abarca todas as atividades que são pensadas e executadas com o intuito de entender como a informação é procurada, transferida e utilizada, dadas as características de cada contexto envolvido. [Gasque \(2008\)](#) possui entendimento similar, sustentando que a busca e o uso da informação estão diretamente vinculadas ao contexto em que o indivíduo está inserido, de maneira que ao se considerar este fator, positiva-se a tendência da obtenção de resultados mais eficazes e aderentes à realidade dos indivíduos e de suas necessidades.

Uma das principais referências acadêmicas sobre este tema, [Wilson \(1999\)](#) afirma que existem três elementos essenciais incluídos no campo de estudo em questão: a necessidade da informação, o processo de busca da informação e o uso da informação. O autor define o comportamento informacional a partir de quatro concepções complementares:

- Comportamento informacional: refere-se ao conjunto de comportamentos humanos observados quanto ao uso de canais e fontes de informação, englobando ainda a sua busca por meio de atitudes passivas ou ativas.
- Comportamento de busca de informação: trata-se de realizar uma ação de busca de informação como efeito da necessidade de alcançar determinado objetivo.
- Comportamento de pesquisa de informação: consiste em um nível do comportamento no qual interações com sistemas de informação de diversas categorias são conduzidas pelas pessoas.
- Comportamento de uso da informação: envolve todas as ações físicas e intelectuais requeridas para que seja incorporado aos conhecimentos prévios de uma pessoa as novas informações obtidas.

Os três elementos indicados por [Wilson \(1999\)](#) estão presentes na maioria dos modelos criados para esquematizar um conjunto de atividades que habilitem a compreensão do comportamento informacional dos indivíduos de uma organização. Esses modelos são

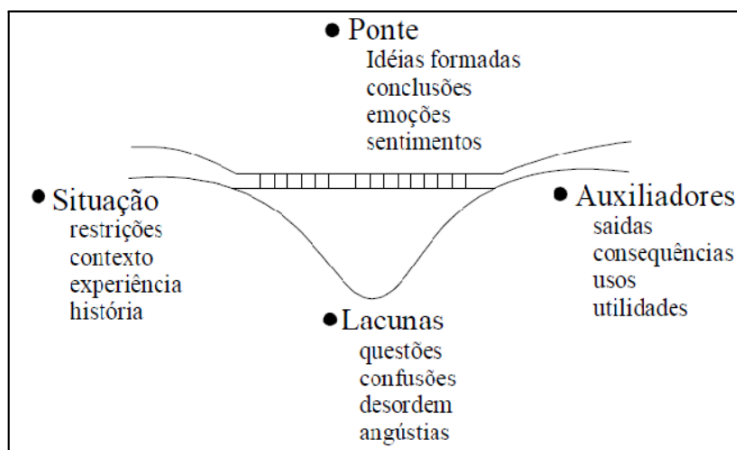
compostos de definições, usualmente apresentadas como diagramas, descrevendo uma atividade para busca de informação, as causas e consequências de sua execução e os relacionamentos desta etapa com outras do fluxo informacional. Entretanto, construir uma inter-relação entre os modelos criados no âmbito dos estudos deste tema não é uma tarefa simples (e até mesmo adequada, em determinadas situações), devido ao fato do conjunto de atividades proposto em cada um deles ser diferente.

Dentre os diversos modelos de comportamento informacional propostos ao longo dos anos, destacam-se o de [Dervin \(1983\)](#), conhecido como *sense-making*, o de [Ellis \(1989\)](#), o de Wilson (1996) e o de Choo (2006), nomeado de integrativo ou multifacetado. É pertinente sublinhar que não é objetivo desta tese apresentar em detalhes as características de cada modelo, mas tão somente caracterizá-los no contexto desta área de estudos, em especial o modelo de Wilson (1996) dado o seu emprego como suporte para alcance de um dos objetivos específicos elencados no Capítulo 1.

O modelo de *sense-making* foi sendo construído por Brenda Dervin desde o início da década de 1970, até que em 1983 a autora propôs uma formulação mais definitiva visando contribuir para os estudos na área de Ciência da Informação. O modelo envolve um grupo de premissas no nível conceitual e métodos avaliativos que visam gerar subsídios para o entendimento sobre como o sentido é construído na realidade das pessoas e como a informação é empregada neste processo, tendo como pano de fundo o atendimento às necessidades individuais ([GONÇALVES, 2012](#)). Sob a ótica de [Rozados \(2003\)](#), o foco do *sense-making* reside na compreensão sobre a forma mais adequada de lidar com a necessidade da informação, a qual está ligada diretamente com a relevância, no sentido de identificar aquilo que tem sentido. A realização de ações para busca de informação são orientadas, portanto, a partir de alguma lacuna na estrutura de conhecimento do indivíduo.

Metaforicamente, conforme explicado por [Bax e Dias \(1997\)](#), o modelo envolve quatro elementos fundamentais analisados sob a peça de um indivíduo atravessando uma ponte: uma situação, uma lacuna, um resultado e uma ponte. Espera-se construir pontes capazes de conectar lacunas à possíveis resoluções com o apoio de auxiliares, considerando também a situação atual envolvida. A Figura 8 expõe esses quatro elementos no contexto da metáfora do *sense-making*. O ponto de partida do processo de busca de informação é uma situação (S) composta por conceitos e ideias influenciadas por restrições de diversas ordens, pelo contexto em questão, bem como pela experiência e história do indivíduo. As lacunas (L) indicam o problema a ser resolvido para remover o impedimento para acesso aos auxiliares pelo indivíduo. Os auxiliares (A) atuam sob o formato de conceitos ou ideias geradas como resultados e representam a consequência da execução do processo de *sense-making*. A ponte (P) é composta por ideias, atitudes e percepções, sendo responsável por interligar a situação e os auxiliares, sobrepondo as lacunas.

Figura 8: Esquema da metáfora do *sense-making*



Fonte: [Bax e Dias \(1997\)](#)

O modelo de David Ellis foi concebido, inicialmente, em 1989, passando por um aperfeiçoamento em 1993. De acordo com [Costa, Silva e Ramalho \(2009\)](#), este modelo de comportamento foca a busca de informação e destina maior importância para os aspectos cognitivos neste processo. Ellis não promoveu a modelagem no formato diagramático, optando por definir um conjunto de etapas gerais, não sucessivas e independentes, fato que possibilita iniciar o processo de busca da informação a partir de qualquer uma dessas etapas. Tal flexibilidade é entendida como uma vantagem deste modelo por [Tabosa e Pinto \(2016\)](#). As etapas do modelo são assim descritas por [Barros \(2008\)](#):

- Inicialização: Indica o começo da busca de informação a partir da percepção de uma necessidade; é o momento onde o indivíduo providencia a construção de uma visão panorâmica sobre o tema de interesse.
- Ligação ou Encadeamento: Abrange as ações do indivíduo para aprofundar a sua busca por meio de ligações entre as primeiras fontes encontradas e referências adicionais.
- Navegação: Ainda se trata de uma análise superficial, pelo indivíduo, das unidades de informação encontradas, podendo ser classificada como uma busca semiestruturada e informal; neste momento, espera-se descobrir itens com maior probabilidade de atender as necessidades originais.
- Diferenciação: Utilizando procedimentos mais rigorosos e critérios como autoria, atualidade e qualidade da fonte, o indivíduo realiza a filtragem e seleção das fontes de informação que mais lhe interessam.
- Monitoramento: Envolve o monitoramento de determinadas fontes de informação, como determinados tipos de periódicos, por exemplo.

- **Extração:** Trata-se da etapa na qual o indivíduo permanece por mais tempo, pois nela acontece o uso efetivo da informação recuperada para atendimento da necessidade originária identificada na inicialização.
- **Verificação:** Nesta etapa o indivíduo promove uma conferência com o intuito de verificar a veracidade das informações extraídas.
- **Finalização:** Indica o momento para que o indivíduo finalize o processo de busca de informação, decidindo por formalizar os resultados ou conduzir novas buscas por publicações mais recentes.

O modelo de Thomas Wilson foi originalmente elaborado em 1981 enfocando os padrões do comportamento informacional das pessoas, passando por revisões abrangentes em 1996 e renomeado para modelo de comportamento informacional complexo, dada sua nova configuração dispondo de interligações com outros domínios. O modelo original preconizava a necessidade humana como ponto de partida para a busca da informação, isto é, os indivíduos começam este processo de procura quando existe o desejo de contemplar alguma demanda específica. Embora o autor ressalte que a busca de informações pode ser realizada em fontes alternativas, não somente em locais tradicionais como bibliotecas, o uso da informação obtida é dependente do sucesso alcançado na busca, sendo previsto o reinício do processo quando a busca não for bem sucedida. Uma das limitações do modelo original é que ele apenas destacava as lacunas na investigação dentro do escopo de um mapa da área, evitando sugerir os fatores vinculados ao comportamento informacional ([WILSON, 1999](#)).

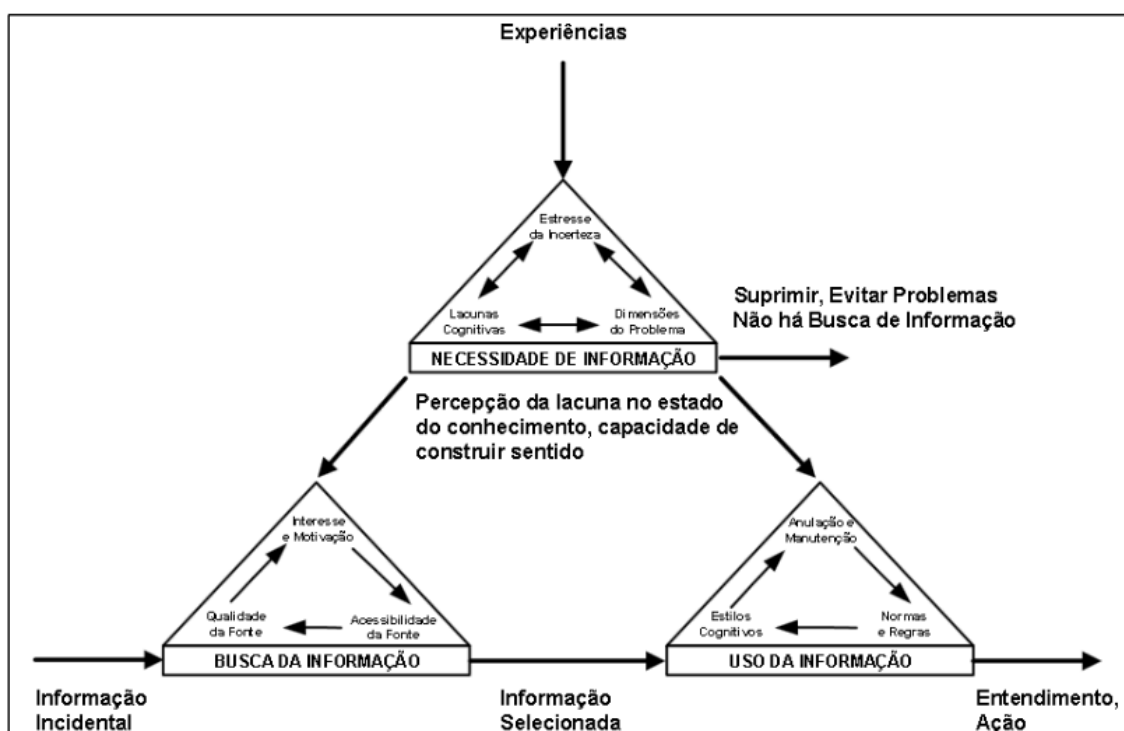
A estrutura central originária foi mantida no novo modelo, inclusive com a continuidade da pessoa em contexto como cerne da necessidade de informação. Foi adicionado um estágio entre a pessoa e a sua consciência da necessidade de informação para representar, assim como no modelo *sense-making*, uma lacuna entre a situação e o uso da informação. Utilizando a teoria do estresse/enfrentamento como suporte, Wilson incluiu no novo modelo um conceito denominado “mecanismo de ativação”, explicando as razões pelas quais determinadas necessidades informacionais não se transformam em processos de busca. Outra evolução refere-se à inclusão de uma fase intermediária, chamada “variáveis intermitentes”, localizada entre a percepção da demanda de informação e a atitude necessária para consegui-la, indicando fatores capazes de obstruir ou desencadear iniciativas de busca de informação. Baseada nas definições propostas pela teoria do risco/recompensa, esta fase contempla análises objetivando mensurar se as contribuições advindas dos esforços de pesquisa suprem os riscos, isto é, se os fatores que beneficiam as iniciativas são mais proeminentes do que aqueles com capacidade de obstrução de tais ações ([WILSON, 1999](#)).

Mais complexo, o segundo modelo de Wilson assume que a busca de informação está relacionada com variados tipos de comportamentos, desde aqueles mais passivos até outros ativos, devendo a GI estar atenta para que sejam oferecidos ambientes adequados a todos esses tipos de perfis. A vinculação do modelo com outras teorias também aumenta a sua carga de complexidade, o que é compensado pelo fato de serem elementos agregadores ao entendimento do comportamento de busca dos indivíduos ([WILSON, 1999](#)).

O modelo multifacetado de Choo foi elaborado para integrar abordagens construídas anteriormente que englobam a necessidade, a busca e o uso da informação. Dentre elas, destacam-se os trabalhos de Wilson e Dervin. Para [Pettigrew, Fidel e Bruce \(2001\)](#), antes do estudo da estrutura do modelo de Choo, é preciso ter ciência de que para este autor a informação é equiparada incorretamente a um recurso localizado em sistemas de informação, documentos ou outros tipos de artefatos, cujas características principais são a constância, a imutabilidade e a dotação de significado a partir da representação no artefato.

Para Choo (2006), a informação deve ser visualizada como o resultado da construção de significado de mensagens por um indivíduo ou grupo de pessoas, e não somente como um objeto. Composto de três fases que contemplam a necessidade, a busca e o uso da informação, cada uma delas englobando fatores cognitivos, afetivos e situacionais, o esquema gráfico do modelo de comportamento informacional integrado de Choo é representado na Figura 9.

Figura 9: Modelo de comportamento informacional de Choo



Fonte: Traduzido de [Choo \(1999\)](#)

Choo (2006) afirma que a fase de necessidade de informação indica o momento em que o indivíduo constrói a percepção de que existe uma lacuna em seu estado atual de conhecimentos, o que pode se tornar, posteriormente, um impulso para iniciar o processo de busca ou uma conclusão de que naquela ocasião é mais adequado suprimir tal percepção. O autor ressalta que a necessidade de informação, muitas vezes, é simbolizada por uma questão ou pergunta direcionadora, a qual não surge plenamente formada, mas é lapidada por processos mentais que a fazem evoluir e encontrar um ponto de resposta possível para o indivíduo.

Neste modelo, a necessidade da informação possui quatro níveis: I) Visceral, onde existe um vago sentimento de insatisfação ou de ausência de conhecimentos, muitas vezes complexo para ser exposto no formato de termos linguísticos; II) Consciente, quando a área de indecisão ou lacuna pode ser descrita com o suporte de informações vagas; III) Formalizado, onde uma descrição da necessidade da informação se mostra racionalizada no formato de questionamento; IV) Adaptado, indicando que o indivíduo está habilitado a adequar a sua pergunta no formato requerido para processamento e entendimento por algum sistema de informação.

A fase de busca de informação engloba as atividades conduzidas pelo indivíduo de maneira a modificar o seu estado atual de conhecimentos. Choo (2006) afirma que alguns comportamentos típicos são observados neste momento, tais como a identificação e seleção de fontes, com destaque para aquelas que ostentam qualidade sob diversos critérios, a articulação mais elaborada de uma pergunta direcionadora, a extração da informação das fontes selecionadas, a avaliação das informações obtidas e, em situações onde o resultado não foi completamente satisfatório para o indivíduo, a modificação da busca.

A terceira fase envolve a escolha daquelas mensagens que se apresentam como mais importantes no espaço da informação construído ou já existente, nos casos em que não houve previamente a busca da informação. Almeja-se, portanto, que o uso da nova informação proporcione mudanças no estado de conhecimento do indivíduo e implique em capacidades de ação que, por sua vez, respondam ao questionamento inicialmente elaborado (CHOO, 2006).

No âmbito da aplicação deste modelo, três propriedades da busca e uso da informação são consideradas. A primeira entende que a busca da informação é construída, sendo que a sua forma e propósito dependem, em grande medida, das estruturas emocionais e cognitivas do indivíduo. A segunda propriedade consiste na consciência de que o uso da informação é situacional, isto é, o contexto para esta utilização é estabelecido a partir da combinação do meio profissional ou social do indivíduo, o ambiente onde este grupo vive ou trabalha, a estrutura dos problemas que o grupo enfrenta no seu cotidiano e as formas empregadas por ele para resolver tais situações. A última propriedade indica a

existência de uma dinâmica para o uso da informação, ressaltando que as três fases do modelo podem ser conduzidas como ciclos recorrentes cujos resultados interagem sem ordem rígida definida (CHOO, 2006).

2.5 Ontologias

Nesta Seção buscou-se conceituar o termo ontologia e promover uma diferenciação entre ontologias de domínio e de fundamentação, indicando suas principais características. O processo de construção de ontologias também é discutido, com destaque para a apresentação da abordagem que suportou o desenvolvimento da ontologia proposta nesta tese.

Uma das principais definições sobre ontologia foi introduzida por [Gruber \(1995, p. 2\)](#) ao afirmar que se trata de um termo advindo da Filosofia com o intuito de indicar a “especificação explícita de uma conceitualização”. Considerando que tal conceitualização refere-se a uma compreensão de mundo simplificada e abstrata a ser representada com um objetivo bem delineado, para o autor é verdadeiro o fato de que qualquer sistema baseado em conhecimento está, direta ou indiretamente, vinculado a alguma conceitualização. No caso dos sistemas de Inteligência Artificial (IA), que necessitam da representação de tudo aquilo que existe em determinado domínio, a representação do conhecimento ocorre sob um formalismo declarativo composto de vários objetos (e suas relações) que, por sua vez, compõem um conjunto de termos com alguma representatividade para o domínio e constituem o universo do discurso. Uma das características das ontologias é possibilitar que esses termos, vistos no formato de classes, relações, funções ou outros objetos, sejam associados a textos capazes de descrever os seus significados e axiomas responsáveis pela restrição da interpretação e uso dos mesmos, de forma que exista um entendimento consistente sobre o vocabulário constituinte da ontologia.

A partir de uma extensão da conceituação proposta por [Gruber \(1995\)](#), [Borst \(1997, p. 12\)](#) define ontologia como a “especificação formal de uma conceitualização compartilhada”. O fato de se tratar de um artefato formal faz com que ele seja passível de leitura por máquinas, motivo pelo qual a área de IA utiliza ontologias em seus sistemas para ser capaz de gerar raciocínios próprios com vistas à resolução de problemas. Ademais, esta formalidade pressupõe a indicação explícita de quais são os conceitos, propriedades, relações e restrições da ontologia, estabelecendo assim o conhecimento consensual advindo do entendimento provido pelos pares. Quanto à conceitualização, trata-se da noção de que a ontologia será, quando finalizada, um modelo abstrato de algum domínio ou fenômeno existente no que conhecemos por mundo real. O autor ressalta, ainda, que esta definição visa explicitar a importância da concordância dos envolvidos com a construção de

uma ontologia quanto à sua conceitualização, de forma que ela seja suficientemente flexível para habilitar o reuso por outras ontologias.

De acordo com [Guarino \(1997\)](#), uma ontologia pode ser compreendida como um artefato composto de um vocabulário vinculado a determinada realidade que, quando utilizado por indivíduos ou sistemas eletrônicos, permite descrever o domínio em questão. Este artefato possui, ainda, a explicitação de fatos aceitos e relacionados com o sentido desejado para aqueles termos constantes do vocabulário, expressos na forma da teoria da lógica de primeira ordem. Espera-se que este artefato represente formalmente os conceitos e as relações entre eles, considerando que previamente houvera a aceitação dos envolvidos quanto ao significado dos termos constituintes do vocabulário. [Uschold e Gruninger \(1996\)](#) apresentam entendimento semelhante, reforçando a importância da característica das ontologias de criarem um vocabulário de termos com a devida especificação dos significados e apreciação conjunta de especialistas, o qual será compartilhado sob a forma de conhecimento.

Segundo [Devedzic \(2002\)](#), o principal objetivo da construção de uma ontologia é a elaboração de um vocabulário para o domínio de determinado problema, no qual conste não somente a definição dos termos usuais para o contexto, mas também quais restrições devem ser aplicadas no processo de modelagem quando da combinação entre tais termos. Outras razões que justificam todo o esforço demandado para a elaboração de uma ontologia são a possibilidade de compartilhar conhecimentos comuns sobre um domínio, explicitar de maneira formal as proposições definidas pelos especialistas e contribuir para a geração de novos conhecimentos tendo como base a reutilização e extensão da ontologia.

Para [Campos \(2019\)](#), um dos principais benefícios das ontologias é possibilitar, por meio da definição de um marco padronizado para interpretação, que os dados sejam interoperáveis e reutilizáveis por plataformas independentes, além de evitar a ocorrência de inconsistências conceituais. As ontologias ainda podem ser diretamente ligadas à ideia de computar o conhecimento, a qual versa sobre a importância de representá-lo em formas digitais executáveis por máquinas ([FRIEDMAN e FLYNN, 2019](#)), demonstrando assim a pertinência deste recurso como organizador das estruturas requeridas pelas máquinas para processar dados e gerar informações válidas que, por sua vez, podem se traduzir em novos conhecimentos úteis para os indivíduos tomarem decisões.

[Almeida e Bax \(2003\)](#) promovem uma visão geral sobre ontologias, mencionando que se trata de um recurso significativo para a área de organização e recuperação do conhecimento. Embora existam diversas definições que caracterizam formalmente as ontologias, os autores adotaram em seu trabalho a conceituação proposta por [Borst \(1997\)](#), dada sua simplicidade e completude. Para eles, cabe à ontologia não somente o agrupamento de termos relacionados com determinado domínio, mas principalmente o

estabelecimento das normas a serem observadas para a combinação entre esses termos, momento em que é criada uma relação baseada na análise e compreensão mútua de especialistas. Desta forma, os usuários estarão aptos a elaborar consultas relacionadas com os conceitos abarcados na ontologia.

Em uma conceituação mais relacionada com a área de Ciência da Informação, [Jacob \(2003\)](#) afirma que uma ontologia se refere ao entendimento semântico, por vários indivíduos, de um domínio específico, normalmente representada em uma estrutura formal baseada em linguagem lógica e contendo as devidas explicações semânticas sobre cada conceito relevante, suas propriedades e também acerca dos relacionamentos entre estes. Tal estruturação formal, quando adequadamente finalizada, permite o processamento de recursos por máquinas capazes de interpretar os conceitos mapeados na ontologia. [Breitman, Felicíssimo e Casanova \(2005\)](#) corroboram este entendimento, complementando que as ontologias possuem grande importância para a interoperabilidade semântica entre distintos sistemas de informação na *web*.

Segundo [Ding e Foo \(2002\)](#), as ontologias trouxeram para o universo da *web* uma possibilidade factível, destinada a melhorar a consistência das informações geradas e processadas em seu ambiente, de maneira que o conhecimento compartilhado pelos diversos sistemas de informação seja lastreado por definições semânticas inteligíveis por humanos e pelas máquinas.

Quanto aos elementos básicos de uma ontologia, [Gruber \(1995\)](#) indicou que quatro itens compõem a sua estrutura, os quais são abordados por [Heinzle \(2011\)](#) com maior detalhamento, conforme exposto na sequência. As **classes** são o elemento principal em uma ontologia, pois representam um conceito existente no domínio que está sendo modelado, são encarregadas de prover a organização do conhecimento envolvido, normalmente são dispostas no formato de uma taxonomia e podem herdar características de outra classe que, por sua vez, passa a ser conhecida como classe-pai. Assim como no modelo Orientado a Objetos, cujo emprego é comum na área de Computação e Sistemas de Informação, a especificação de uma classe na ontologia representa um conjunto de itens com atributos e comportamentos iguais, formando assim o conceito que define um objeto único. As **relações** entre os conceitos modelados na ontologia constituem o seu segundo elemento básico, habilitando compreender qual é o tipo de interação existente entre aqueles conceitos envolvidos em uma relação particular. Os **axiomas** são o terceiro elemento de uma ontologia, sendo utilizados para formalizar as regras que devem ser cumpridas pelos outros elementos, de forma a restringir a interpretação e o emprego dos conceitos modelados. O último elemento consiste nas **instâncias**, as quais servem para representar determinados objetos da ontologia, isto é, alguns exemplares específicos de algumas de suas classes.

Embora diferentes propostas de classificações das ontologias tenham sido indicadas por variados autores ao longo dos últimos anos, existem quatro tipificações mais comuns que são adequadas para complementar o entendimento teórico requerido nesta tese. Essa classificação se origina dos estudos de [Guarino \(1997\)](#), [Guizzardi \(2005\)](#) e [Haav e Lubi \(2001\)](#). Em uma "ontologia de fundamentação ou de alto nível" o objetivo é formalizar a descrição de classes gerais sem dependência direta com algum domínio, abordando conceitos abrangentes como tempo, objeto, espaço e matéria. Por outro lado, na "ontologia de domínio" busca-se explorar um domínio específico para descrever os seus conceitos e considerando que, normalmente, aqueles que fazem parte da estrutura da ontologia são próprios do domínio. Uma "ontologia de tarefa" visa contribuir para a resolução de problemas a partir da compreensão do significado de conceitos especializados de determinada atividade ou tarefa, porém de modo independente de um domínio. Já uma "ontologia de aplicação" exprime dependência direta com algum domínio e organiza as conceituações requeridas de maneira que o conhecimento envolto em determinada aplicação seja modelado.

[Keet \(2020\)](#) ressalta que a definição de uma ontologia de domínio pode envolver abordagens diferentes em função da dimensão analisada. Sob o aspecto semântico, tem-se o levantamento das entidades e suas relações, conjuntamente à representação formal dos significados destes elementos. Tendo o aspecto pragmático como base, itens práticos da construção de ontologias deverão ser pontuados, tais como o nível de conhecimento do domínio investigado, os métodos de projeto que suportaram o desenvolvimento dos artefatos e os critérios de raciocínio. [Carvalho, Almeida e Guizzardi \(2014\)](#) complementam afirmando que este tipo de ontologia precisa ser elaborada como uma ontologia de referência, visando descrever clara e precisamente as características da realidade do domínio em questão ao capturar e formalizar uma conceitualização compartilhada do mesmo.

[Guizzardi \(2007\)](#) defende, ainda, que a construção de ontologias de domínio requer o suporte de uma ontologia de fundamentação apropriada, responsável por indicar os possíveis estados do mundo conforme os compromissos ontológicos fundamentais. Ele argumenta que ontologias de fundamentação fornecem os conceitos ontologicamente mais básicos e universais, os quais servem como base para a construção de ontologias mais específicas, como as de domínio. O alinhamento da ontologia de domínio com a de fundamentação permitiria, de acordo com o autor, a especificação de um subconjunto desses estados conforme sua admissibilidade dentro do escopo de um domínio particular.

Para [Guarino \(1998\)](#), pelo fato de serem artefatos de representação de termos ou entidades genéricas e sem dependência direta com algum domínio específico, as ontologias de fundamentação habilitam os especialistas a construir modelos conceituais com elevado

grau de abstração e estão aptas a colaborar como meio para a resolução de problemas relacionados com a interoperabilidade entre sistemas. [Smith e Brochhausen \(2010\)](#) acrescentam que, ao utilizar uma ontologia de alto nível para abstrair aqueles conceitos mais gerais, torna-se plausível harmonizá-los com os conceitos mais específicos delimitados no escopo da ontologia de domínio, promovendo a integração de conhecimentos heterogêneos e suportando a consecução da interoperabilidade semântica. Para [Lopes Júnior et al. \(2023\)](#), o fato das entidades de diferentes ontologias de domínio se especializarem em um mesmo conceito da ontologia de fundamentação permite a integração semântica de dados entre ontologias de diferentes domínios.

2.5.1 Unified Foundational Ontology (UFO)

De acordo com [Guizzardi, Falbo e Guizzardi \(2008\)](#) e [Guizzardi et al. \(2015\)](#), a UFO (*Unified Foundational Ontology*) é uma ontologia de fundamentação elaborada com o agrupamento de teorias das áreas de ontologia formal, linguística, lógica filosófica e psicologia cognitiva, sendo composta de pequenas teorias relacionadas, especialmente, com aspectos essenciais da modelagem conceitual. Com o intuito de direcionar o seu escopo conforme as possibilidades de utilização, a UFO é categorizada em três estratos. A UFO-A, também denominada ontologia de objetos, foi empregada nesta tese e foca na promoção de maior semântica na modelagem conceitual desenvolvida para determinado domínio de conhecimento, envolvendo aspectos sobre tipos e suas instâncias, objetos e propriedades particulares, relações entre os elementos e regras. Como é uma ontologia de *endurants*, existe a preservação temporal da identidade das entidades, cuja existência pode ser independente. A UFO-B, denominada ontologia de eventos, propõe a existência de elementos que acontecem no tempo ao invés de persistirem nele, como eventos e estados, denominando-os como *perdurants*. Já a UFO-C é uma ontologia fundamentada em UFO-A e UFO-B que busca estruturar em objetos e eventos aqueles conceitos sociais encontrados no contexto de entidades sociais.

Com o propósito de habilitar a reutilização dos recursos da UFO em ontologias operacionais, a gUFO³ (*gentle* UFO) foi especificada como uma versão que satisfaz a demanda por artefatos ontológicos acessíveis diretamente por aplicações computacionais. Uma de suas principais vantagens reside na possibilidade de aproveitamento por instanciação e/ou especialização dos elementos definidos em uma ontologia de referência construída com a UFO, tais como classes, propriedades de objetos e propriedades de dados. Dada a amplitude da UFO, fator que pode, eventualmente, gerar dificuldades de utilização por projetistas não totalmente familiarizados com suas definições, a especificação da gUFO foi concebida para implementar somente uma parte do estrato da UFO-A

³ <https://nemo-ufes.github.io/gufo/>

(englobando os conceitos de Objetos, Aspectos e Situações, além da diferenciação entre Indivíduos e Tipos) e outra parte do estrato da UFO-B (definição de Eventos e a noção de participação neles). Merece apontamento o fato da manutenção pela gUFO das duas taxonomias presentes originalmente na UFO, uma dispondo de classes cujas instâncias são indivíduos (*gufo:Object* ou *gufo:Event*) e outra onde as instâncias criadas são tipos/universais (*gufo:Kind* ou *gufo:Phase*, por exemplo) ([ALMEIDA et al., 2019](#)).

2.5.2 Construção e representação de ontologias

[Gruber \(1995\)](#) estabelece que a construção de ontologias precisa ser entendida no contexto de um projeto formal, no qual várias decisões são tomadas com o intuito de fazer com que o artefato produzido esteja apto a uma avaliação criteriosa da sua qualidade e aderência. O autor propõe a análise de cinco critérios para avaliação do projeto de ontologias, os quais objetivam compartilhar conhecimentos e possibilitar a interoperação de programas que compartilham uma conceitualização:

- Clareza: Analisa se as definições da ontologia são completas e objetivas ao ponto de comunicarem efetivamente o significado dos seus termos.
- Coerência: Indica se as inferências produzidas a partir da ontologia estão logicamente de acordo com suas definições.
- Extensibilidade: Analisa a capacidade da ontologia ser extensível para novos usos e permitir a incorporação, nestes casos, de novos termos sem impactar as definições originais.
- Viés de codificação mínimo: Avalia se a ontologia foi construída de maneira a possibilitar sua implementação em diferentes sistemas e estilos de representação.
- Comprometimento ontológico mínimo: Verifica se a ontologia define somente aqueles termos suficientes para comunicar o conhecimento do domínio em questão.

[Guarino \(1997\)](#) faz um contraponto à proposta de [Gruber \(1995\)](#) para criação de ontologias mais genéricas e extensíveis ao afirmar que o processo de construção precisa considerar a real complexidade e especificidade das características dos objetos que compõem a conceitualização, abarcando potenciais relações futuras entre eles, ante uma representação que explore tão somente o estado atual dos objetos. Esta conceitualização mais complexa e completa pode auxiliar, inclusive, na garantia de que ambiguidades serão evitadas, dado o maior nível de detalhamento dos objetos representados.

As vantagens do emprego de ontologias são resumidas por [Guizzardi \(2000\)](#) sob três aspectos: 1) Amplia a comunicação entre todos os interessados no domínio investigado ao possibilitar a obtenção de consenso sobre os termos envolvidos a partir de entendimentos

comuns que, por sua vez, promovem a geração de novos conhecimentos; II) Contribui para a eliminação das inconsistências e contradições do conhecimento representado, dada a sua formalização não ambígua em uma estrutura que pode ser verificada e validada; III) Constituem um vocabulário consensual com alto nível de abstração sobre os termos abarcados em um domínio específico, fomentando, assim, a reutilização da ontologia em outros contextos.

Por outro lado, autores como [Noy e McGuinness \(2001\)](#), [Rector et al. \(2004\)](#) e [Sales e Guizzardi \(2015\)](#) ressaltam que a construção de ontologias é um processo complexo, demandando envolvimento direto dos pesquisadores, especialistas de domínio e potenciais usuários para que os erros de desenvolvimento venham a ser tratados ainda na etapa de projeto. Dentre os tipos de erros mais comuns, destacam-se a utilização de relações que não se mostram genuinamente ontológicas, a sobrecarga de relações “é-um”, as definições imprecisas de classes e suas relações, o emprego de “classes miscelâneas”, a utilização de herança múltipla e a criação de polissemias. Uma lista de erros mais extensa e com maior nível de detalhamento pode ser encontrada no trabalho de [Orlando \(2017\)](#).

A representação de uma ontologia pode ser feita com o suporte de diversas tecnologias, dentre as quais merecem indicação o RDF (*Resource Description Framework*), RDFS (*RDF Schema*) e OWL (*Ontology Web Language*). Segundo [Antonou et al. \(2012\)](#), um modelo baseado em RDF realiza a representação das informações por meio de triplas contendo sujeito (recurso), predicado (propriedade) e valor (declaração), podendo ser mapeado como um grafo no qual os sujeitos e os valores são os vértices e as propriedades consistem nos elos de ligação entre estes. O RDF é independente de domínio e não possui sintaxe própria, mas é possível mapear os seus elementos para elementos XML (*eXtensible Markup Language*), por exemplo, utilizando-os como *tags* em conjunto com um RDFS, pois este permite a definição do vocabulário do RDF. O RDFS dispõe de conceitos ontológicos para estruturar as informações no formato de um esquema, tais como classes, hierarquia, propriedade, domínios e restrições.

No contexto da Web Semântica, o W3C (*World Wide Web Consortium*) construiu a linguagem OWL para auxiliar na elaboração de ontologias que requerem maior expressividade, mantendo como sua base o RDF e o RDFS. Os elementos fundamentais em uma ontologia representada com OWL são as entidades, as quais podem ser denominadas como classe, indivíduo, *object property* ou *data property*, e os axiomas, responsáveis por indicar o que é entendido como verdade em determinado domínio. A linguagem *Turtle* é uma das sintaxes que podem ser escolhidas para expressar os dados no modelo requerido pelo RDF.

2.5.3 Integração semântica de dados baseada em ontologia

A ampliação do número de bases de dados mantidas dentro e fora do ambiente das organizações e o uso de sistemas de informações providos por diferentes fornecedores, ocorridos no contexto da explosão do volume de dados eletrônicos processados com os mais diversos propósitos, fez com que a demanda por uso desses repositórios se desse de forma integrada, enriquecendo o conjunto de dados original com dados oriundos de fontes complementares. Tal processo de integração envolve uma complexidade muitas vezes alta demais para ser suportada apenas por simples estratégias de manipulação dos repositórios, pois estes são independentes em termos de funcionamento e, normalmente, apresentam formatos próprios em sua estrutura ([BITTAR et al., 2018](#); [CRUZ et al., 2019](#)).

[Farinelli, Almeida e Melo \(2013\)](#) citam outras dificuldades envolvidas com o processo de integração, como a heterogeneidade dos dados quanto à sua natureza e conteúdo e a vasta disponibilidade de recursos tecnológicos não necessariamente padronizados que produzem dados continuamente, tais como equipamentos que funcionam no escopo da Internet das Coisas. No âmbito da Administração Pública, a adoção de soluções que atuam no sentido de promover a interoperabilidade dos dados está diretamente relacionada com a qualidade dos serviços prestados aos cidadãos, aumentando assim a percepção de importância da definição de padrões que mitiguem problemas de integração entre sistemas de informação e seus dados.

As diferenças semânticas existentes em conjuntos de dados que necessitam de integração computacional podem ser resolvidas com o apoio de ontologias, conforme [Pessanha e Bax \(2016\)](#) e [Uschold e Jasper \(1999\)](#), pois elas são capazes de construir uma especificação, sem ambiguidade, dos termos utilizados em cada domínio, contextualizando e significando os dados. Para [Guarino \(1998\)](#), as ontologias contribuem para a interoperabilidade semântica de dados com a definição do significado dos termos e das relações entre esses termos dentro de um domínio específico, explicitando, assim, como eles devem ser interpretados por pessoas e máquinas. A designação de um significado para cada termo pode ser restringida com a aplicação de axiomas que, enquanto formalizam o aceite de determinadas interpretações, rejeitam outras.

De acordo com [Mendonça e Zaidan \(2019\)](#), as ontologias são importantes para integrar dados heterogêneos, atuando como uma metodologia para a especificação precisa de termos do domínio envolvido e das relações entre eles, permitindo entendimento global livre de ambiguidades. A abordagem ontológica amplia as possibilidades de construção informacional pelas próprias máquinas, pois os dados passam a dispor de uma estrutura uniforme que os torna passíveis de integração e compartilhamento. [Castro e Simionato \(2021\)](#) complementam afirmando que, em um projeto de integração semântica, as ontologias simbolizam um esquema global no qual será baseado o mapeamento entre as diferentes

fontes de dados, conceituando os termos do domínio relacionado e aumentando a sua capacidade semântica.

A partir do momento em que uma ontologia harmoniza o entendimento sobre os termos oriundos de diferentes fontes e que estão envolvidos no processo de integração, têm-se um recurso no qual estará especificada a semântica dos termos que compõem este novo domínio, formando um vocabulário global que possibilita o mapeamento das fontes de dados ([AUCELI, BERARDI e KOZIEVITCH, 2019](#)).

[Britto, Ruy e Azevedo \(2020\)](#) reforçam a validade do uso de ontologias como referências semânticas na integração de dados que, mesmo obtidos em variadas origens e dotados de significados e formatos diferentes, são mapeados para uma estrutura uniforme que padroniza o acesso por aplicações heterogêneas. Outra vantagem do emprego de ontologias na integração de fontes de dados é apontada por [Almeida, Ribeiro e Maculan \(2019\)](#), indicando que, a partir do momento em que elas são caracterizadas como estruturas de representação do conhecimento consensual, contribuem para que a linguagem informal existente na prática dos domínios dê lugar à sustentação formal requerida pelo contexto digital.

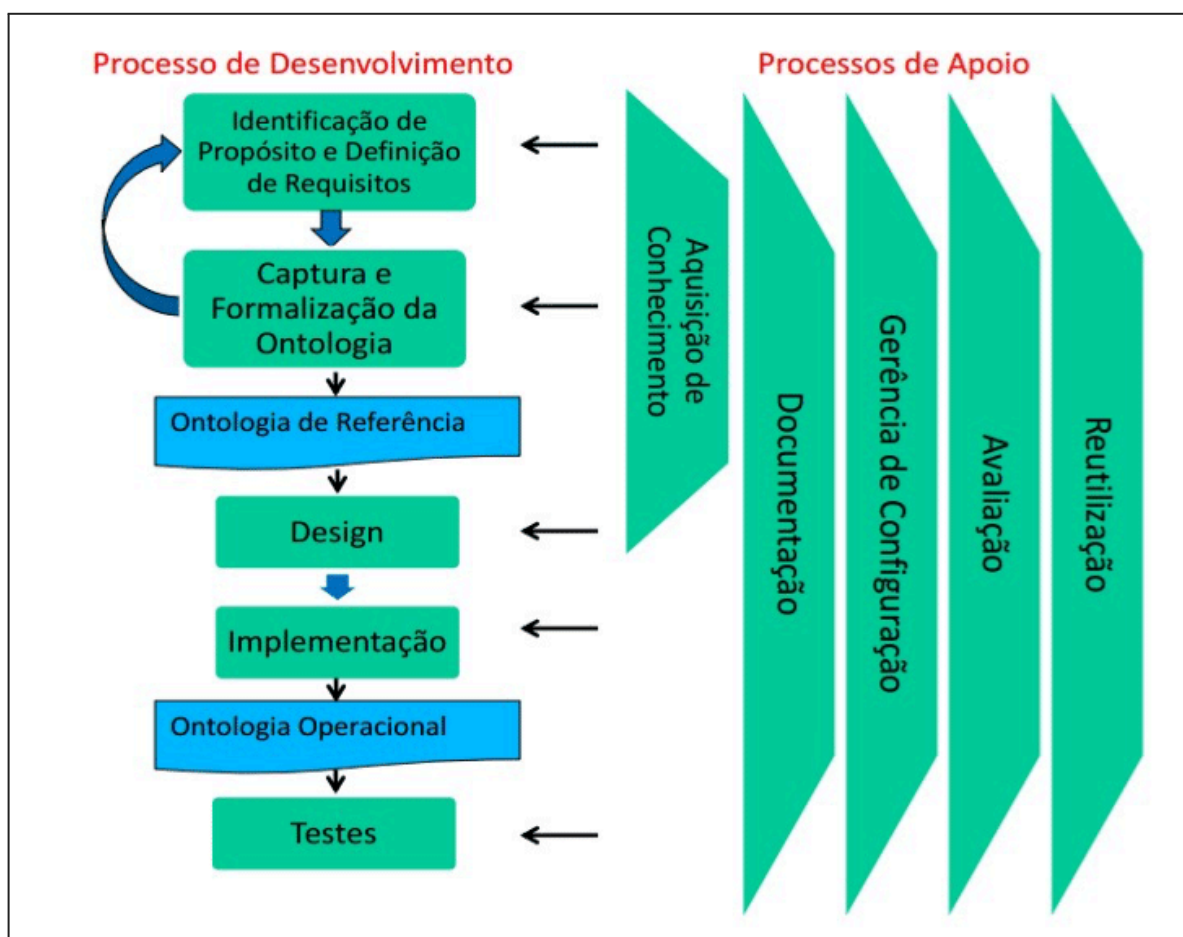
2.5.4 Abordagem de engenharia de ontologias (SABiO)

O processo de criação de ontologias com qualidade não é trivial, pelo contrário, os desafios para que o conhecimento seja organizado sob esta abordagem são inúmeros e, por vezes, complexos. Assim como ocorre no desenvolvimento de produtos de software, é necessária a aplicação de diversos métodos, técnicas e ferramentas que suportem todas as etapas e atividades do processo, uma sub-área de estudos denominada Engenharia de Ontologias ([FALBO, 2014](#)). Partindo deste princípio, a utilização de metodologias que auxiliem na estruturação do projeto de criação da ontologia pode ser um fator benéfico não apenas para os indivíduos envolvidos, mas também para o produto final a ser produzido.

Neste sentido, [Falbo \(2014\)](#) introduziu a abordagem denominada *Systematic Approach for Building Ontologies* (SABiO) como uma metodologia que permite conceber, implementar e validar ontologias de domínio. Trata-se de um método sistematizado para a elaboração de ontologias utilizado em diversos estudos acadêmicos ([DETONI et al., 2018](#); [GUARNIER et al., 2020](#); [MADDALENA E BAIÃO, 2021](#); [SILVEIRA et al., 2021](#); [SOUZA, FALBO e VIJAYKUMAR, 2017](#)) que, por sua vez, atestaram a capacidade desta metodologia como organizadora do processo de desenvolvimento de ontologias. Sua estrutura é relativamente objetiva e adaptável no que tange às atividades e diretrizes para o seu emprego. Uma das recomendações da SABiO indica a importância de se analisar os conceitos e relações especificados em uma ontologia de referência à luz de uma ontologia de fundamentação, como a UFO.

Todo o detalhamento a seguir das características da metodologia foi baseado em [Falbo \(2014\)](#), motivo pelo qual esta referência não é repetidamente mencionada. De acordo com a Figura 10, é possível notar que a SABiO é composta de um processo macro (**Desenvolvimento**) constituído de fases próprias e vários processos de **Apoio**. É importante ressaltar que não se trata de uma metodologia rígida em relação à sequência de execução destas fases, sendo recomendado que todo o projeto de desenvolvimento da ontologia privilegie um ciclo incremental e iterativo. Outra menção válida requer o entendimento que o foco da SABiO é contribuir para a construção de ontologias de domínio, seja como uma ontologia de referência (vista sob a ótica de um tipo particular de modelo conceitual) ou como uma ontologia operacional (uma versão que permite a leitura e interpretação da ontologia por máquinas).

Figura 10: Processos e fases da Metodologia SABiO



Fonte: Adaptado e traduzido de [Falbo \(2014\)](#)

Durante a fase inicial, denominada “Identificação de Propósito e Definição de Requisitos”, espera-se elicitare e estudar os requisitos envolvidos para aumentar o entendimento sobre questões mais abrangentes da ontologia, principalmente pela

elaboração das questões de competência (QC) que guiarão o restante das atividades. É fundamental deixar bem explicitado o que a ontologia precisa estar apta a responder quando do seu uso, demonstrando então o quanto seu propósito é coerente com a realidade. Nessa fase também ocorre a análise da demanda por modularização da ontologia, em especial quando o domínio envolvido possui tamanha complexidade que sua divisão pode ser expressa no formato de sub-ontologias, assim contribuindo para o aumento da facilidade no desenvolvimento e manutenção do projeto.

A fase seguinte contida na metodologia, nomeada como “Captura e Formalização da Ontologia”, visa capturar toda a base conceitual da ontologia e formalizá-la, mantendo sempre a conexão com as questões de competências delimitadas anteriormente. Cabe ressaltar que a SABiO prevê a participação dos indivíduos no projeto de desenvolvimento da ontologia a partir da sua atuação em diversos papéis, como especialista em domínio, usuário de ontologia, engenheiro de ontologia, designer/projetista, programador e testador.

No caso da fase em questão, assume-se a relevância do especialista em domínio como principal contribuidor para a identificação dos conceitos e relacionamentos a serem organizados pela ontologia de referência, além de uma atuação em parceria estreita com os engenheiros de ontologia. A organização de tais conceitos e relações pode ser feita em uma taxonomia a fim de que não ocorra interpretações indevidas dos termos levantados. Sugere-se o emprego de linguagem natural incrementada com exemplos para a descrição dos conceitos primitivos da ontologia, evitando possíveis inconsistências na compreensão e ambiguidades na interpretação. O reuso de outras ontologias, seja no seu todo ou em parte, deve merecer atenção nessa fase. De forma que o processo de comunicação entre especialista de domínio e engenheiro de ontologia seja facilitado, é sugerida a elaboração de um modelo gráfico que permita o estabelecimento de consenso entre os significados para os elementos da ontologia. A SABiO recomenda como linguagem de representação da ontologia o perfil da OntoUML ([Guizzardi, 2005](#)), pois este incorpora os elementos centrais da ontologia de fundamentação UFO.

Caso o projeto da ontologia demande seu uso em aplicações computacionais, a sequência de fases da SABiO precisa ser iniciada visando a elaboração de uma ontologia operacional. Neste caso, busca-se, a partir da ontologia de referência construída na fase anterior, projetar uma ontologia operacional de acordo com determinada linguagem que possa ser lida e interpretada por máquinas, como a OWL (*Ontology Web Language*). Ferramentas computacionais, como o *OntoUML Lightweight Editor*⁴, permitem que os modelos expressos em OntoUML sejam automaticamente convertidos em OWL, contribuindo para a atuação do designer/projetista de ontologia. Nesta fase também é importante levantar e discutir a respeito dos requisitos não funcionais que devem ser

⁴ <https://github.com/nemo-ufes/ontouml-lightweight-editor>

cumpridos a partir da utilização da plataforma de implementação escolhida, considerando aspectos tanto da arquitetura quanto da tecnologia em si. Encerrado o projeto da ontologia operacional, passa-se à fase de implementação, momento em que a modelagem é transformada em um produto de software de acordo com a plataforma escolhida.

A fase de testes é a última indicada na metodologia, buscando realizar a validação e a verificação dos produtos construídos nas fases anteriores, isto é, a especificação da ontologia, a ontologia de referência e a operacional. Recomenda-se que o conjunto de testes a ser executado possua direta ligação com as questões de competência que nortearam todo o projeto da ontologia. É esperado que um caso de teste possua três elementos: I) uma consulta no ambiente informacional implementado visando responder a determinada questão de competência; II) a entrada de dados instanciados de acordo com a parte a ser testada da ontologia; III) a saída esperada conforme a instanciação feita. Os casos de teste também podem analisar questões vinculadas com os requisitos não funcionais do projeto, especialmente aqueles que venham interferir na interação do ambiente informacional com os usuários da ontologia.

Quanto aos cinco processos de apoio definidos na SABiO, cada um possui sua devida importância para o projeto da ontologia. O processo de aquisição de conhecimento perpassa toda a metodologia em si, pois ele pode acontecer mesmo após a ontologia ter sido testada e estar em pleno uso. A participação dos especialistas de domínio é imprescindível, constituindo-se em principal fonte deste processo, aliada à utilização de fontes documentais, como materiais bibliográficos e outros modelos de referência. O processo de reutilização está diretamente ligado à elaboração da ontologia, conforme mencionado anteriormente, permitindo que ocorra o aproveitamento de artefatos ontológicos, parcial ou totalmente. O processo de documentação visa registrar formalmente os resultados das atividades que compõem a metodologia, padronizando com o apoio de *templates* os artefatos preenchidos. O processo de gerenciamento da configuração envolve todos os artefatos construídos durante o projeto da ontologia, sejam eles documentos textuais, diagramas, esquemas gráficos ou código-fonte. Por fim, o processo de avaliação é compreendido sob os pontos de vista da verificação e da validação da ontologia, nos quais espera-se ter uma visão dos resultados a partir de revisões técnicas nos produtos desenvolvidos.

2.6 Ciência de Dados

A área de Ciência de Dados abrange os processos, modelos e tecnologias que atuam com os dados ao longo de todo o seu ciclo de vida, desde a produção, armazenamento, transformação, análise e descarte. Trata-se de uma área que não é

fechada em si mesma, isto é, sua composição advém do que outras ciências elaboram e disponibilizam para manipulações necessárias aos dados, seja no formato de teorias, tecnologias, processos ou procedimentos ([AMARAL, 2016](#)). No caso desta tese, o emprego de recursos da Ciência de Dados, notadamente aqueles vinculados com a subárea de Aprendizado de Máquina, está diretamente relacionado aos métodos para a consecução de alguns objetivos específicos propostos, culminando com um entendimento mais completo acerca dos construtos analisados.

Na sequência, é promovido um entendimento sobre algumas das etapas do ciclo de vida dos dados proposto por [Amaral \(2016\)](#), com foco nas atividades vinculadas com o processo de análise dos dados.

2.6.1 Produção e armazenamento dos dados

Os dados digitais são produzidos de variadas formas, sejam aqueles que advém da interação humana com algum dispositivo ou aqueles gerados pelos próprios sistemas e aplicações computacionais. Uma infinidade de dispositivos são utilizados neste processo, como os teclados, mouses, telas de toque, leitores de códigos, *scanners*, identificadores por radiofrequência, câmeras de vídeo, dentre outros. O processamento da folha de pagamento de uma empresa ou a execução de modelos de aprendizado de máquina são exemplos de rotinas capazes de, por si só, produzirem dados digitais ([MAGRANI, 2018](#)).

Com a recente evolução dos *smartphones*, ficou em evidência outra categoria de dispositivos produtores de dados, os sensores. Em um *smartphone* existem diversos tipos de sensores, como o acelerômetro, GPS (*Global Positioning System*), giroscópio e outros relacionados com a comunicação do aparelho, como o *Bluetooth*, WiFi e NFC. Esta popularização permitiu que fossem construídos sensores para os mais variados fins e áreas de negócios. O volume de dados produzido pelos sensores normalmente é alto, pois alguns deles chegam a executar suas rotinas em menos de segundos, repetindo o mesmo processo continuamente ([FACCIONI FILHO, 2016](#)).

Conforme apontado por [Freitas e Mendes \(2017\)](#), há de se destacar o enorme volume de dados produzido atualmente, fato que traz consigo, pelo menos, duas óticas de análise. De acordo com [Choi, Wallace e Wang \(2018\)](#), [Hasan et al. \(2020\)](#) e [Lies \(2019\)](#), para determinados segmentos de negócio, tais como o marketing digital, tecnologia da informação e aplicações financeiras, a quantidade de dados disponíveis muitas vezes é diretamente proporcional à melhoria da atuação empresarial, isto é, são organizações/instituições que demandam grandes volumes de dados que são analisados com o auxílio de modernas tecnologias para que os produtos e serviços ofertados estejam devidamente alinhados com as demandas e expectativas do mercado consumidor. Neste

contexto, a etapa de produção de dados passa a ter maior significância, englobando variados meios para sua realização.

[Bittencourt et al. \(2018\)](#) e [Hariri, Fredericks e Bowers \(2019\)](#) analisam esta questão por um prisma diferente, no qual é enfatizado que a produção e o manuseio de grandes volumes de dados não são considerados como um processo trivial, pois tal demanda faz com que o número de atividades do ciclo de gestão informacional cresça tanto em quantidade como em complexidade. Tarefas como coleta dos dados, padronização de valores, cumprimento dos requisitos de armazenamento, limpeza de dados inconsistentes, integração de diferentes fontes e processamento reforçam a necessidade de um adequado planejamento sobre quais dados efetivamente precisam fazer parte do ecossistema informacional da organização.

A produção dos dados digitais está diretamente relacionada com a definição dos meios para o seu armazenamento, pois a execução de atividades futuras no ciclo informacional depende, em grande parte, das características desses meios. Os recursos de armazenamento não envolvem somente o espaço utilizado pelos dados, mas também qual o formato e a estrutura necessária para os registros/valores, a localização física ou virtual dos repositórios, além de aspectos de segurança e confidencialidade. Amaral (2016) promove um histórico dos vários modelos de armazenamento construídos e utilizados desde a década de 1960 para permitir que as aplicações computacionais fossem capazes de incluir e alterar os dados produzidos pelos usuários: hierárquico, em rede, relacional, orientado a objetos e NoSQL. O autor realiza, ainda, uma importante distinção entre dados estruturados, aqueles que possuem uma configuração pré-determinada, e não estruturados, que não dispõem de qualquer tipo de organização e, em alguns casos, nem mesmo metadados.

[Avazpour, Grundy e Zhu \(2019\)](#) e [Saggi e Jain \(2018\)](#) sintetizam esta questão ao afirmar que, dada a ampla variedade de recursos disponíveis atualmente para o armazenamento de dados brutos digitais, nenhum modelo pode ser considerado genericamente como o melhor, pois é importante que seja buscada aquela solução mais adequada a cada situação. O mais comum é que em um mesmo ambiente coexistam diversos meios habilitados não apenas ao armazenamento em si, mas também para contribuir na integração dos dados persistidos.

2.6.2 Preparação e transformação dos dados

De acordo com [Oreski, Pihir e Konecki \(2017\)](#), a partir do momento em que os dados brutos estão devidamente armazenados na estrutura selecionada, é comum a realização de atividades que busquem prepará-los para utilização em contextos que demandam, por exemplo, a transformação de valores em um formato específico ou a criação de novos dados a partir daqueles existentes. Objetiva-se, nesta etapa, executar tarefas sobre os

dados de maneira que, ao final, seja possível obter um conjunto de dados adequado à elaboração de análises diversas. [Ramos et al. \(2020\)](#) afirmam que a etapa de preparação e transformação dos dados pode consumir mais da metade do tempo de um projeto de ciência de dados, sendo que a qualidade dos seus resultados é fundamental para que o conhecimento produzido no processo seja confiável e válido.

No escopo da área de Mineração de Dados, conforme apontado por [Costa et al. \(2013\)](#), a totalidade de atividades conduzidas para transformar os dados e prepará-los até a construção de modelos é denominada comumente de etapa de pré-processamento. Uma das principais justificativas para sua importância dentro de um projeto dessa área reside no fato de que, atualmente, existe grande disponibilidade de bases que podem ser integradas para a construção de um único conjunto de dados, cada uma delas com sua própria padronização. Dessa forma, é preciso garantir, por meio da execução de atividades específicas, que os dados estejam padronizados antes de iniciar os procedimentos da etapa de análise. Por exemplo, a medida de determinado item pode estar armazenada em centímetros em uma base de dados e em milímetros em outra, indicando a necessidade de padronização para o caso de tais dados serem integrados.

[García et al. \(2007\)](#) entendem que a etapa de pré-processamento dos dados pode ser condensada em cinco atividades: I) Discretização dos dados, que visa transformar valores numéricos contínuos em categorias; II) Criação de novos atributos a partir daqueles originais (derivação), permitindo a obtenção de determinada informação de maneira mais eficaz; III) Seleção de atributos, que objetiva reduzir o conjunto de atributos para otimizar a análise dos dados; IV) Criação de tabelas de sumarização, para que os dados de bancos relacionais possam ser agrupados em apenas uma estrutura de tabela; V) Transformação do formato dos dados, quando é requerido um formato específico e diferente daquele utilizado pelos dados.

Segundo [Batista \(2003\)](#), é necessário considerar que a etapa de pré-processamento dos dados demanda, na maioria das vezes, grande conhecimento sobre o domínio dos dados, além do fato de que dificilmente a primeira extração dos dados de uma base já proverá um conjunto que não apresente erros, inconsistências ou possibilidades de melhorias. Neste sentido, o autor apresenta uma relação mais ampla de atividades que, quando devidamente conduzidas, contribuem para a melhora da qualidade dos dados antes do seu processamento em modelos de análise. São elas: identificação de inconsistências (quando um mesmo dado é representado por valores diferentes), verificação de integridade (envolvendo a análise das aceitáveis relações entre os atributos), identificação de atributos duplicados e redundantes (quando o mesmo dado está disponível em atributos diferentes), valores *defaults* (representam falta de informação ao invés de informação valorosa), tratamento de valores desconhecidos (ocorre quando determinado atributo não possui valor

em alguns registros, requerendo o seu preenchimento com alguma coerência, como a média dos outros valores, por exemplo), identificação de valores extremos (podem ocasionar distorções na aplicação de testes estatísticos), tratamento de desbalanceamento de classes (trazem riscos de enviesar a análise dos dados), seleção de atributos (para que as análises considerem um número menor de atributos com a mesma capacidade de indução de conhecimento do que se todo o conjunto fosse utilizado) e construção de atributos (produção de novos atributos com maior capacidade para descrever determinado conceito).

A realização integral desse conjunto de atividades não é necessária em todos os casos, sobretudo quando os dados a serem analisados já oferecem adequada formatação. É necessário avaliar qual das atividades faz mais sentido em cada situação, pois o seu tempo de execução pode impactar negativamente no cronograma do projeto. A correta definição de quais atividades da etapa de pré-processamento serão executadas contribui para que o conjunto de dados esteja íntegro, aderente ao proposto e disponível em tempo adequado ([SILVA, 2018](#)).

Outra importante atividade dessa etapa consiste na integração de diferentes fontes de dados, as quais podem ser bases de dados em Sistemas Gerenciadores de Banco de Dados, planilhas eletrônicas, arquivos no formato textual, dentre outros. [Batista et al. \(2018\)](#) ressaltam que, quando a integração de múltiplas fontes é conduzida de maneira alinhada com os objetivos de aquisição de conhecimento definidos, o conjunto de dados construído tende a ser mais rico em termos de informações e proporcionar a resolução de problemas reais. Contudo, [Silva \(2014\)](#) alerta que tais integrações demandam maior nível de atenção por parte dos analistas envolvidos, buscando evitar redundância de dados, sobreposição de registros, dependências entre os atributos e conflitos entre valores preenchidos.

2.6.3 Análise de dados

No escopo da etapa de análise de dados existe uma ampla variedade de atividades que podem ser executadas sob a organização de determinadas metodologias, bem como um grande número de tecnologias informacionais construídas especialmente para promover análises nos dados digitais. Essa Subseção apresenta, brevemente, uma conceituação dos recursos de análise dos dados empregados neste trabalho, os quais estão baseados nas técnicas de Aprendizado de Máquina (AM).

[Cerri e Carvalho \(2019\)](#) conceituam o Aprendizado de Máquina como uma sub-área da Inteligência Artificial que atua na criação de algoritmos capazes de, a partir da experiência adquirida com a análise de dados de treinamento, aprender a executar completamente diversos tipos de tarefas, como o agrupamento e a classificação. O objetivo dessas tarefas pode ser descritivo, quando os dados são analisados para descrever determinado fenômeno, ou preditivo, onde busca-se prever a ocorrência do fenômeno

estudado. Para [Faceli et al. \(2011\)](#), os recursos de Aprendizado de Máquina contribuem diretamente com a construção de hipóteses ou funções a partir da análise de experiências anteriores, servindo de meio para obtenção de definições generalizadas sobre o conjunto de dados investigado.

[Monard e Baranauskas \(2003\)](#) compreendem o Aprendizado de Máquina como uma área que busca desenvolver e aplicar técnicas computacionais baseadas no aprendizado de uma realidade, além de construir sistemas dotados de capacidade para obter novos conhecimentos automaticamente. [Mitchell \(1997\)](#) possui linha de entendimento similar, afirmando que o Aprendizado de Máquina permite o desenvolvimento de novas habilidades e apoia, sob variados aspectos e formas, a organização do conhecimento existente. [Batista \(2003\)](#) adiciona à essa conceituação inicial que a capacidade de aprendizado das máquinas modernas é a base para o seu comportamento inteligente, o que [Sousa \(2020\)](#) compreende como a habilidade da máquina aprender sobre um assunto específico com o menor nível possível de intervenção humana ao longo do processo.

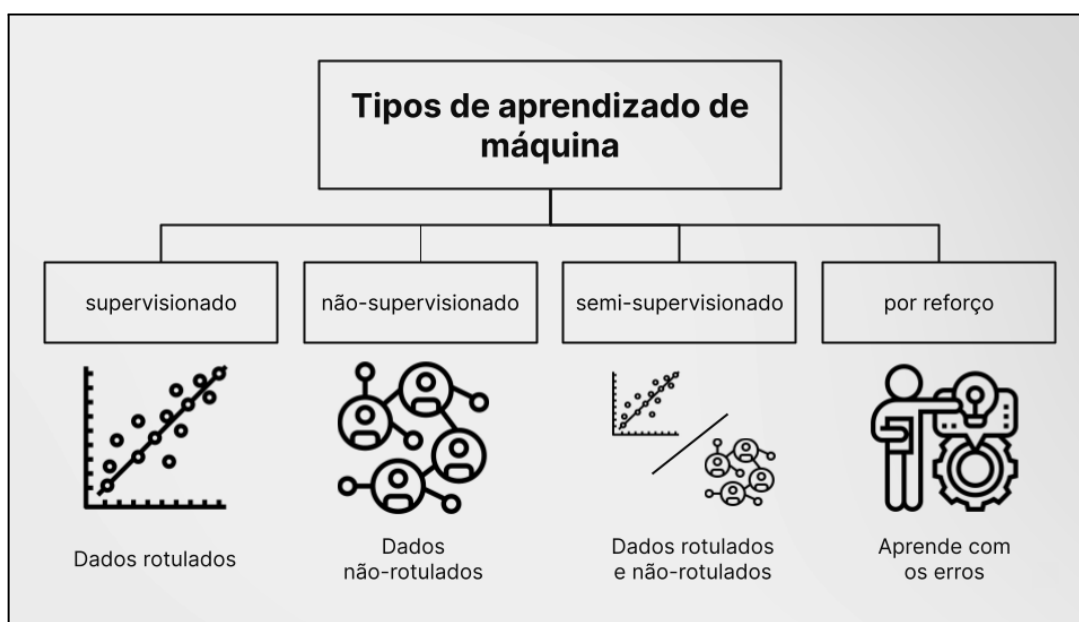
[Naqa e Murphy \(2015\)](#) relacionam o Aprendizado de Máquina com o emprego de algoritmos computacionais, cujo propósito é emular a inteligência humana para obter aprendizado a partir da análise do ambiente contido em amplos conjuntos de dados, permitindo que os mais variados campos de atuação se beneficiem de suas aplicações, como visão computacional, reconhecimento de padrões, medicina e engenharia espacial. [Helm et al. \(2020\)](#) indicam uma conceituação semelhante ao afirmarem que as técnicas de Aprendizado de Máquina são baseadas na aprendizagem mediada por experiências, próxima à realizada pela inteligência humana, permitindo não somente a identificação de padrões nos dados analisados, mas principalmente o seu entendimento e a geração de conclusões mais precisas.

De acordo com [Carmo et al. \(2018\)](#), as técnicas mais populares de Aprendizado de Máquina estão relacionadas com o conceito de indução, abordagem na qual as situações particulares promovem a elaboração de conclusões generalistas, as quais são interpretadas como sendo o aprendizado em si. Conforme os autores, trata-se do conceito mais aplicado e estudado atualmente na área de Aprendizado de Máquina. Ademais, é plenamente factível que o pesquisador construa modelos mais aprimorados com seus dados, pois a geração do aprendizado não está necessariamente ligada apenas a esse escopo, mas sim envolvida em um entendimento mais abrangente acerca do problema trabalhado.

O Aprendizado de Máquina é categorizado por [Ayodele \(2010\)](#) em quatro tipos: aprendizado supervisionado, aprendizado não supervisionado, aprendizado semi-supervisionado e aprendizado por reforço. A Figura 11 representa graficamente esta categorização do Aprendizado de Máquina, em conjunto com a característica principal dos dados em cada um dos quatro tipos. É conveniente mencionar que nesta tese são utilizadas

técnicas do aprendizado não supervisionado para cumprimento do primeiro objetivo específico (“Caracterizar o perfil dos candidatos em processos seletivos da assistência estudantil e construir um modelo de clusterização para identificação de grupos entre os participantes”).

Figura 11: Tipos de Aprendizado de Máquina



Fonte: Raphaell (2021)

Conforme explicam [Monard e Baranauskas \(2003\)](#), no aprendizado supervisionado os dados analisados na etapa de treinamento dos modelos já possuem os rótulos conhecidos, isto é, os resultados finais a serem encontrados pelos algoritmos estão disponíveis no próprio conjunto de dados. Dessa forma, os algoritmos conseguem entender e, principalmente, aprender como aquele resultado foi encontrado. Após a etapa de treinamento e aprendizado, o algoritmo está apto a realizar a predição em um conjunto de dados sem indicação dos rótulos. Nos casos em que o valor do rótulo é discreto, tem-se uma tarefa de classificação, ao contrário das tarefas de regressão, identificada em situações onde o valor do rótulo é contínuo.

[Jordan e Mitchell \(2015\)](#) contribuem nessa caracterização dos métodos de aprendizado supervisionado ao afirmarem que o processo de predição em dados com rótulos até então desconhecidos é viabilizada com a aplicação de métodos de probabilidade sobre esses dados, visando encontrar qual é a saída y para cada entrada x . Tais métodos podem ser baseados em várias formas de mapeamento, como as máquinas de vetores de suporte, árvores de decisão, florestas de decisão, redes neurais ou regressão logística.

O aprendizado não supervisionado é conceituado por [Fisher, Pazzani e Langley \(2014\)](#) como aquele em que a definição do rótulo dos dados não é previamente conhecida dentro do escopo do conjunto de dados utilizado nas tarefas de treinamento dos algoritmos, demandando a consecução de técnicas para analisar os dados e, baseado em observações diversas, especificar um possível rótulo a cada instância, permitindo assim agrupá-las conforme esta rotulagem. [Monard e Baranauskas \(2003\)](#) complementam este entendimento ao destacarem que o aprendizado não supervisionado investiga os dados com o apoio de algoritmos que objetivam formar eventuais agrupamentos das instâncias analisadas (também chamados de *clusters*). Concluída esta etapa de elaboração dos agrupamentos, é possível que as características de cada um deles sejam investigadas para que colaborem com maior entendimento sobre o problema que se pretende resolver.

[Norman \(2021\)](#) ressalta que, considerando a ausência de rótulos previamente definidos no conjunto de dados, os algoritmos utilizados em problemas de aprendizado não supervisionado têm liberdade para encontrarem suas próprias conclusões ao compararem os dados consigo mesmos, auxiliando na descoberta da estrutura geral dos dados ou da sua distribuição. [Araújo Filho \(2022\)](#) destaca a capacidade dos algoritmos conseguirem reconhecer padrões nos dados a partir dos atributos em comum identificados entre as instâncias analisadas, formando os agrupamentos conforme a similaridade desses atributos e possibilitando que novas instâncias sejam vinculadas a um agrupamento existente quando da sua investigação.

Para [Gibson, Rogers e Zhu \(2013\)](#), o aprendizado semi-supervisionado é uma categoria intermediária entre o aprendizado supervisionado e o não supervisionado, pois nele o conjunto de dados possui, ao mesmo tempo, dados com e sem rótulos. Para [Guo, Wang e Li \(2021\)](#), o grande atrativo deste tipo de aprendizado reside no fato de que não necessariamente é preciso descartar dados não rotulados no processo de construção dos modelos de Aprendizado de Máquina. Os dados rotulados do conjunto são utilizados nas tarefas de treinamento de um modelo que, posteriormente, fará a predição dos rótulos das instâncias não rotuladas, gerando instâncias denominadas como pseudo-rotuladas. Na sequência, os dados originalmente rotulados são combinados com aqueles pseudo-rotulados para que o modelo de classificação seja novamente treinado em um conjunto de dados com maior quantidade de instâncias, perfazendo, ao final, um processo puro de aprendizado supervisionado.

[Wang e Taylor \(2017\)](#) especificam o aprendizado por reforço como uma abordagem que visa aumentar e melhorar o aprendizado dos algoritmos a partir da aquisição de novas experiências obtidas com a aplicação da técnica de tentativa e erro. Os resultados são alcançados com a realização de treinamentos com um agente que conduz interações com o meio para incrementar cumulativamente as recompensas e, deste modo, reforçar o que fora

aprendido até então. [Jordan e Mitchell \(2015\)](#) esclarecem que, de maneira oposta ao que ocorre no aprendizado supervisionado, as técnicas de aprendizado por reforço utilizam os dados de treinamento somente para apontar se uma ação específica está correta ou não, ficando a cargo do pesquisador encontrar a solução para aquelas situações identificadas como incorretas.

2.6.3.1 Agrupamento de instâncias

Uma das principais tarefas dos algoritmos de aprendizado não supervisionado consiste em agrupar as instâncias de um conjunto de dados conforme as características em comum que elas apresentam, parâmetro também conhecido como medida de proximidade. A partir da definição de elementos similares, os algoritmos conseguem delimitar os componentes de cada agrupamento sem (ou com baixa) supervisão humana, razão pela qual o conjunto de técnicas que trabalham com dados sem rótulos pré definidos para treinamento de modelos é denominado como aprendizado não supervisionado ([AGGARWAL, 2018](#)).

Conforme apontado por [Hastie, Tibshirani e Friedman \(2009\)](#), dois fatores influenciam diretamente o processo de agrupamento. O primeiro deles é a medida de proximidade, responsável por definir como será calculada a similaridade entre duas instâncias. A escolha pela medida de proximidade mais adequada para cada situação depende, em grande parte, dos tipos de atributos encontrados no conjunto de dados. Por exemplo, para dados contínuos, é possível utilizar as medidas de similaridade conhecidas como Euclidiana, Manhattan ou Pearson, enquanto para dados binários o coeficiente de Jaccard pode ser escolhido. O segundo fator que impacta nos resultados do processo em questão é o método de agrupamento, cujas estratégias podem ser categorizadas como particional ou hierárquico. Enquanto os algoritmos da categoria particional dividem o conjunto de dados em um compartimento simples de k grupos, aqueles hierárquicos constroem uma sequência de compartimentos aninhados, organizando as instâncias analisadas em grupos e subgrupos.

[Marcacini \(2014\)](#), [Miraftabzadeh et al. \(2023\)](#) e [Sutramiani et al. \(2024\)](#) indicam o algoritmo *K-means* como uma possível escolha nos diversos cenários que demandam agrupamento de dados sob a categoria particional. Trata-se de um método de aprendizado de máquina não supervisionado para conjuntos de dados numéricos que busca criar k agrupamentos (parâmetro que pode ser informado pelo próprio usuário) baseando-se na aproximação de dados semelhantes após o registro da distância (Euclidiana, por exemplo) entre os elementos. Para cada agrupamento, é calculado um valor denominado “centróide”, representando a média da distância entre os objetos que estão localizados mais próximos a

este quando comparado com outros centróides. Portanto, tem-se o centróide como um objeto que, de maneira geral, representa todos os elementos de um agrupamento.

[Tan et al. \(2019\)](#) ressaltam que o particionamento gerado pelo algoritmo, isto é, se os elementos dentro de um agrupamento são realmente semelhantes uns aos outros e, ao mesmo tempo, dissimilares dos elementos dos outros agrupamentos, pode ter sua qualidade avaliada a partir da coleta da métrica *Silhouette Score*. Trata-se de uma pontuação que representa a eficácia do algoritmo no que tange a separar corretamente os elementos em cada agrupamento e também separar os agrupamentos em si. A pontuação, também chamada de coeficiente de silhueta, pode variar de -1 a 1: resultados mais próximos de 1 indicam que os elementos alocados em cada agrupamento estão distantes dos elementos dos outros agrupamentos, havendo, portanto, uma fronteira bem definida entre os agrupamentos; resultados próximos de 0 ou inferiores indicam alta proximidade no limite de decisão para alocar os elementos entre dois agrupamentos ou, até mesmo, atribuição de elementos em agrupamentos errados.

De acordo com [Napoleon e Pavalakodi \(2011\)](#), o processo de agrupamento pode sofrer dificuldades na etapa de visualização dos resultados quando envolver conjuntos de dados com alta dimensionalidade, isto é, contendo grande quantidade de atributos. Esta característica pode tornar complexa a tarefa de definir corretamente as distâncias entre cada elemento analisado, acarretando em agrupamentos compostos de elementos dissimilares.

Uma das mais eficientes técnicas para redução da dimensionalidade em conjuntos de dados é a *Principal Component Analysis* (PCA), capaz de transformar dados de alta dimensão em um espaço de baixa dimensão composto por um conjunto menor de atributos que, por sua vez, condensam estatisticamente em poucos componentes a informação contida em diversos atributos originais, permitindo assim sua visualização em gráficos com duas ou três dimensões. Mesmo com a redução para um espaço de dimensão inferior, normalmente cerca de 90% da variância da informação do conjunto de dados original é mantida no que a técnica denomina como componentes principais. O primeiro componente principal irá armazenar a maior parte das informações dos recursos originais, o segundo armazenará a próxima maior quantidade de informações, e assim sucessivamente. Cada componente principal é independente e contém informações únicas. Feito isto, o algoritmo de agrupamento pode utilizar, na sua análise, os componentes principais encontrados ao invés de todos os atributos originais do conjunto de dados, mantendo quase o mesmo nível de variância da informação.

O *K-means* é um algoritmo com implementação relativamente simples e de execução iterativa, conforme observado por [Oliveira \(2018\)](#). No Algoritmo 1 tem-se a sua representação no formato de pseudocódigo. São passados dois parâmetros de entrada para o algoritmo, a quantidade de agrupamentos a serem construídos e o conjunto de dados com

os seus atributos e elementos. A primeira tarefa consiste na escolha aleatória de quais serão os k centróides dos agrupamentos definidos inicialmente, isto é, os elementos representantes de cada agrupamento. Em seguida, o algoritmo entra em um laço de repetição que executa duas tarefas, atribuindo cada objeto analisado ao agrupamento em que o centróide está mais próximo e recalculando os centróides para atualizar os representantes dos agrupamentos. O encerramento do laço ocorre quando o número máximo de iterações for executado (caso essa definição tenha sido registrada no início do procedimento) ou quando nenhuma mudança nos centróides é mais necessária, indicando que as mudanças de objetos entre os agrupamentos não conseguem minimizar a função objetiva.

Algoritmo 1: Pseudocódigo do algoritmo *K-means*

- 1: **procedimento** K-means(k : agrupamentos, D : conjunto de dados)
- 2: escolher aleatoriamente k objetos de D como centróides dos agrupamentos iniciais;
- 3: **repetir**
- 4: (re) atribuir cada objeto ao agrupamento do centróide mais próximo;
- 5: atualizar o centróide de cada agrupamento e minimizar a distância intra agrupamento;
- 6: **até que** agrupamentos sem alterações;
- 7: **fim do procedimento**

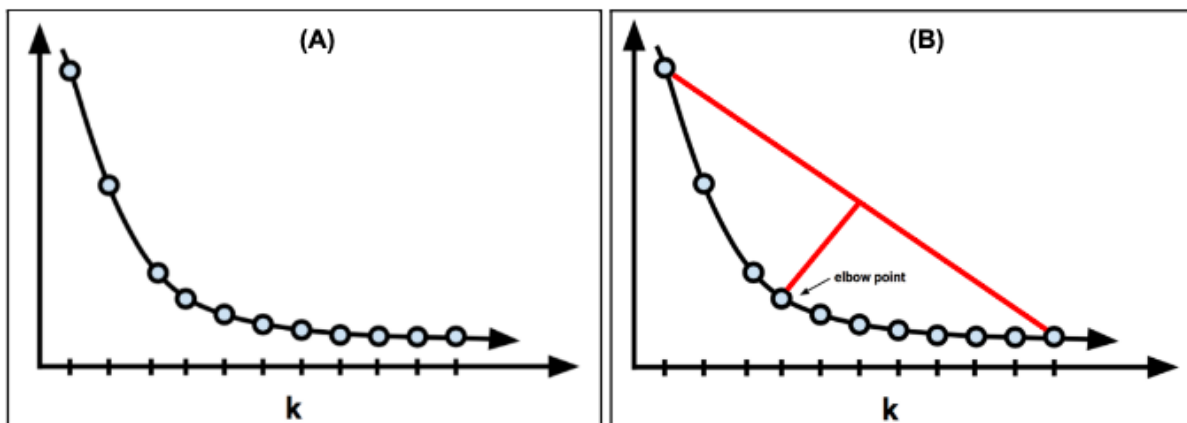
Fonte: Adaptado de [Oliveira \(2018\)](#).

Uma possível dificuldade de utilização do *K-means* reside na complexidade de estipular, sem subsídios prévios, a quantidade de agrupamentos a ser construída pelo algoritmo. [Marcacini \(2014\)](#) afirma que esta deve ser uma preocupação real, pois a definição inadequada ou com baixa cobertura de testes e análises pode gerar grande variabilidade dos resultados. Para minimizar os efeitos de tal definição, o autor sugere que o algoritmo seja executado diversas vezes, cada qual com uma inicialização diferente do parâmetro k , permitindo a escolha daquela solução com o menor valor de erro.

[Nogueira \(2022\)](#) também adverte para essa situação, pois a inadequada definição do parâmetro k pode impactar a qualidade dos agrupamentos encontrados. Caso ele seja muito pequeno, a homogeneidade dos elementos intra agrupamento não será ampla; caso seja muito grande, o reconhecimento de padrões nos dados ficará prejudicado. O autor aponta a utilização de uma estratégia conhecida como Método *Elbow*, baseada na geração do gráfico de um fator representativo da dispersão dos dados, como a soma dos erros quadrados, por exemplo, no eixo y , para os valores correspondentes de k no eixo x . Conforme observa-se na Figura 12(A), cada ponto no gráfico indica o centróide daquele agrupamento. A partir de determinado momento, percebe-se uma tendência à constância nos valores do eixo y , a

despeito do contínuo incremento nos valores de k . Quando uma reta é traçada para interligar os centróides localizados nas extremidades, similar ao apresentado na Figura 12(B), identifica-se o ponto inicial de estabilização dos valores de y no elemento mais distante da reta, sendo este considerado como o mais adequado valor de k para execução do algoritmo *K-means*.

Figura 12: Exemplo de aplicação do Método *Elbow*



Fonte: [Nogueira \(2022\)](#)

Na próxima Seção, algumas publicações que exploram temáticas similares às conduzidas nesta tese são caracterizadas e discutidas à luz dos objetivos aqui definidos, fornecendo assim maior conhecimento acerca das relações entre os três construtos investigados.

2.7 Avaliação de políticas públicas

Antes de tratar especificamente do processo de avaliação de uma política pública, é necessário compreender, sob o ponto de vista teórico, qual o significado e a aplicação deste conceito. Neste sentido, [Lopes e Amaral \(2008\)](#) afirmam que as políticas públicas abrangem os planos, ações e metas propostos pelos governos em suas diversas esferas, como nacional, estadual ou municipal, com o intuito de promover o bem-estar de toda a sociedade em consonância com o interesse público. Dado que as demandas dos grupos aos quais os indivíduos pertencem não necessariamente são as mesmas, e os recursos disponíveis pelo poder público para atender aos pleitos são escassos, as políticas que prosperam neste ambiente precisam, sob pena de impacto negativo em seus resultados, ser de interesse da maior parte da sociedade envolvida, incrementando o bem-estar daqueles participantes direta ou indiretamente com suas ações.

De acordo com [Saravia e Ferrarezi \(2007\)](#), o conceito de política pública é passível de análise mediante duas perspectivas: abstrata e operacional. No caso da abstrata, ela é entendida como um conjunto de decisões públicas tomadas com o propósito de continuidade do equilíbrio social ou para construir um desequilíbrio apto a alterar a realidade vivida pela sociedade. Sob o enfoque operacional, uma política pública se traduz como um sistema formalizado de recursos, objetivos, metas e ações que, estabelecido um período específico de tempo de atuação, mantenha ou modifique, preferencialmente para melhor, diversos aspectos da vida social de um grupo de indivíduos.

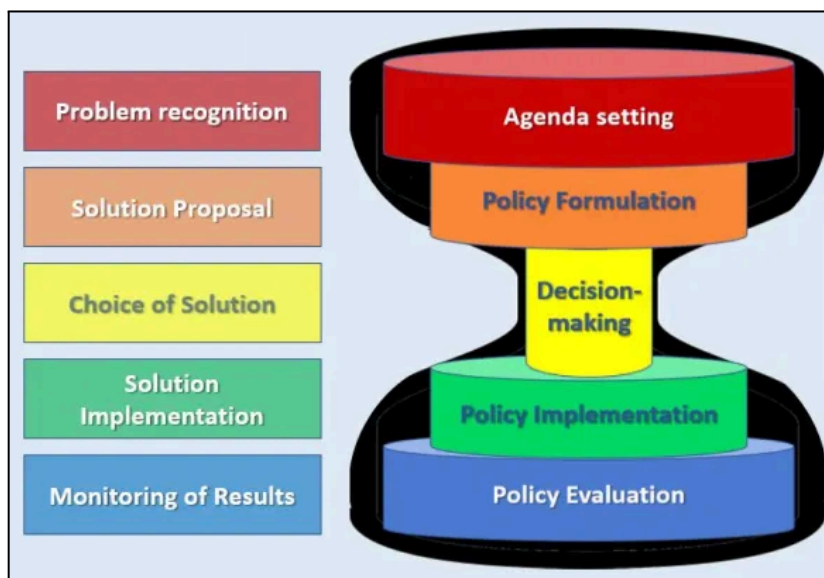
Para [Vásquez e Delaplace \(2011\)](#) uma política pública pode ser identificada em contextos onde o poder governamental utiliza seus próprios recursos na promoção do bem-estar da sociedade com a execução de ações racionais e formalmente delineadas. Dessa forma, a política pública seria o instrumento para a solução de problemas que afetam a normalidade do bem-estar dos indivíduos. Os autores ressaltam a necessidade de as atividades constantes no processo de elaboração e execução da política pública serem guiadas pela racionalidade dos envolvidos, fazendo com que o resultado alcançado seja adequado aos objetivos propostos, sem redução ou incremento exagerados do escopo de atuação da política pública.

Entendimento similar é apresentado por [Secchi \(2014\)](#), cuja conceituação de política pública aponta para uma linha de ação constituída a partir do objetivo de atender a sociedade em problemas do seu cotidiano. É fundamental, na visão deste autor, que a política seja estruturada como uma resposta intencional a estes problemas, seja ativa ou passiva, evitando desvios inadequados de sua abrangência e, ao mesmo tempo, atendendo a uma relevância coletiva.

No que diz respeito às características gerais das políticas públicas, [Viana \(1996\)](#) afirma que elas precisam ter bem clara a definição do público-alvo; apresentar-se no formato de medidas concretas sobre o contexto em que está inserida; indicar os critérios e meios pelos quais os recursos serão alocados e distribuídos; façam parte de um conjunto maior de ações capitaneadas pelo governo; fixem objetivos e metas a serem atingidas em momentos próprios de medição. Para [Saravia e Ferrarezi \(2007\)](#), as políticas públicas mantêm, mutuamente, quatro propriedades: I) institucional, indicando que uma autoridade constituída formal e legalmente chancelou sua elaboração; II) decisória, tendo sido constituída para representar uma decisão governamental que visa contribuir para a solução de problemas e/ou atendimento de necessidades; III) comportamental, delimitando eventuais ações ou inações conforme cada contexto apresentado; IV) causal, representando a consciência do impacto da execução de suas proposições na sociedade.

[Howlett, Ramesh e Perl \(2009\)](#) relacionam cinco fases como componentes básicos do ciclo para construção de políticas públicas, expostas na Figura 13.

Figura 13: Fases do ciclo de construção de políticas públicas



Fonte: [Howlett, Ramesh e Perl \(2009\)](#)

A primeira fase é nomeada como “Formação da agenda” e visa compreender os problemas da sociedade para, na sequência, identificar as questões prioritárias a serem tratadas pelo governo. As linhas de ação para resolver essas questões são delineadas na segunda fase (“Formulação de políticas”), tendo como resultado uma análise de viabilidade das propostas analisadas sob os aspectos financeiro, legal, técnico e político. Na terceira fase (“Tomada de decisões”) são especificados o prazo e os recursos de determinada ação, preferencialmente formalizados como atos da administração pública. A quarta fase (“Implementação”) envolve a transformação do planejamento da política pública em execução de atividades, isto é, colocar em prática o que fora previsto legalmente. A “Avaliação” é a última fase do ciclo e consiste no mapeamento de indicadores e na realização de suas medições para que os resultados possam ser compreendidos sob o ponto de vista do alcance dos objetivos propostos pela política pública. Os autores ressaltam que não se trata de um ciclo rígido quanto à sequência de execução, estando aberta a possibilidade de interações entre as atividades de fases diferentes.

O caráter dinâmico do ciclo de formulação de políticas públicas é ressaltado por [Rua \(2009\)](#) ao afirmar que não se trata de um processo linear e fragmentado, mas dotado de características que requerem interações contínuas entre os envolvidos independentemente da fase em questão. Esta interatividade é requerida, ainda, em função do contexto no qual as políticas públicas são construídas, abarcando variadas relações de poder e estranhamento entre as demandas da sociedade e da classe política. Considerando que nesta tese o foco recai sobre a avaliação do PNAES enquanto política pública, apenas os elementos da quinta fase do ciclo de vida serão abordados a seguir.

O primeiro aspecto a ser considerado no âmbito da avaliação de políticas públicas, de acordo com [Pereira \(2017\)](#), indica a importância da definição de sistemas de monitoramento e avaliação que permitam aferir a sua efetividade de acordo com a proposta de criação. Tal processo é capaz de fornecer subsídios, em momentos oportunos, para a identificação de eventuais desvios na realização da política, habilitando os gestores a elaborarem e executarem medidas corretivas em tempo hábil. Ademais, a organização governamental e a sociedade civil podem se beneficiar da aprendizagem promovida pela análise sobre o alcance das metas traçadas para cada ação da política. [Ham e Hill \(1993\)](#) complementam reiterando que os resultados estratégicos de uma política estão diretamente ligados à consecução, com sucesso, das atividades de monitoramento e avaliação.

[Baptista e Rezende \(2015\)](#) abordam o caráter vertical do processo de avaliação, mencionando que não se trata de somente executar um conjunto de atividades ao final do período previsto para determinada política pública, mas de promover momentos ao longo de todo o processo para averiguar a conformidade dos resultados obtidos em cada etapa. Essas análises não ficam, portanto, restritas à época de encerramento da política, mas se mostram relevantes justamente por formarem subsídios intermediários que auxiliam na tomada de decisão e na condução das atividades específicas das fases do ciclo de formulação de políticas públicas.

A fase de avaliação é classificada por [Arretche \(1998\)](#) em três modalidades para que fique mais claro o objeto de análise em cada uma delas. A “avaliação política” envolve exame do contexto prévio à construção da política em si, buscando fundamentar sua razão de existência sob o ponto de vista político do processo decisório. A “análise de políticas públicas” compreende a engenharia institucional necessária para que o processo de implementação aconteça com previsibilidade quanto aos fatores facilitadores e eventuais obstáculos, embora não preveja preocupação acerca dos prováveis resultados. A última modalidade é nomeada como “avaliação de políticas públicas”, onde busca-se investigar, com o suporte de métodos formais de pesquisa, a relação de causalidade entre a política e os resultados alcançados, bem como estimar os impactos promovidos por ela no público envolvido e também na sociedade. Neste sentido, [Khandker, Koolwal e Samad \(2010\)](#) destacam a importância da utilização de métodos preferencialmente aplicados em outros estudos ou com o devido lastro teórico para medir se as intervenções geradas pela política pública impactaram os resultados de interesse.

[Secchi \(2014\)](#) alerta que a concepção dos processos de avaliação de políticas públicas precisa contemplar eventuais complicações trazidas pela multicausalidade, isto é, quando diversos fatores contribuem, em variados níveis, para influenciar os resultados observados, não sendo completamente possível afirmar que as alterações ocorridas na sociedade decorreram, exclusivamente, por conta da ação da política pública. Ademais, a

identificação de tais fatores e a compreensão do papel específico de cada um deles sobre os resultados não é algo trivial, motivo pelo qual [Baker \(2000\)](#) propõe, para assegurar o devido rigor metodológico, que a avaliação de impacto de políticas públicas aconteça com a denominação de um contrafactual equivalente a um contexto no qual a política não existe. Assim, a estimação do impacto em um cenário com baixa probabilidade de influência de causas não controladas traria uma medida mais confiável da real participação da política pública na transformação da sociedade envolvida.

2.8 Trabalhos correlatos

Os estudos analisados nesta Seção contribuíram, cada um de forma específica, para a elaboração desta tese, sendo apresentadas suas características básicas e comentados os resultados obtidos. Cabe ressaltar que todo o referencial teórico mencionado até aqui aportou, de certa maneira, subsídios particulares e úteis para a construção e desenvolvimento de noções mais amplas sobre os construtos envolvidos, as formas de integração e avaliação dos mesmos.

[Barbosa, Ferreira e Rezende \(2020\)](#) objetivaram avaliar e acompanhar os resultados do PNAES no contexto da Universidade Federal de Alfenas (UNIFAL-MG), construindo, para isso, dois indicadores relacionados com as taxas de evasão e retenção de estudantes de cursos de graduação matriculados entre 2014 e 2016. Dado que o PNAES se propõe a contribuir com esses aspectos, o estudo identificou a sua efetividade quanto aos dois quesitos na maioria das turmas avaliadas, com destaque para aquelas cujo ingresso ocorreu em 2014, uma vez que houve menos evasão e retenção entre os estudantes beneficiados pelo Programa. Diferentemente do proposto neste trabalho, o estudo analisado admitiu como estudantes assistidos pelo PNAES aqueles que receberam qualquer tipo de benefício assistencial, seja auxílio permanência, alimentação, creche, participação em eventos ou atividades pedagógicas.

Enfocando apenas a relação entre a concessão de auxílio permanência e evasão escolar, estudo similar havia sido conduzido por [Saccaro, França e Jacinto \(2016\)](#), no qual foi encontrado resultado positivo entre os construtos, demonstrando que a taxa de evasão era cerca de 6% mais baixa entre aqueles estudantes que fizeram jus ao recebimento do benefício. Outro resultado indicado pelos autores aponta que esta melhora na taxa de evasão aconteceu independentemente do tipo de benefício recebido pelo estudante, comprovando a influência positiva do recebimento de auxílios previstos no PNAES.

[Betzek \(2015\)](#) analisou os indicadores de permanência e evasão de uma IFES compreendendo o período de 2008 a 2013, encontrando forte relação positiva entre o recebimento de auxílios socioeconômicos ao longo do tempo e a permanência do estudante

na instituição. De acordo com a pesquisa, aqueles estudantes que receberam qualquer tipo de auxílio por apenas um semestre letivo apresentaram taxa de evasão média de 35%; estudantes que receberam auxílio por dois semestres tiveram taxa de evasão próxima a 17%; quando o benefício se estendeu a quatro semestres, a taxa chegou a 9%. Outra análise feita pela autora constatou que a quantidade de auxílios diferentes concedidos a um mesmo estudante possui relação diretamente proporcional com o aumento de suas chances de continuar matriculado no curso, reduzindo assim os níveis gerais de evasão escolar.

A conclusão do estudo de [Machado, Oliveira e Freitas \(2020\)](#) indica que a distribuição de auxílios socioeconômicos ocasionou impacto positivo no desempenho acadêmico dos estudantes de uma IFES brasileira. Foram analisados os dados demográficos e acadêmicos com referência ao ano de 2015, permitindo caracterizar os estudantes tanto sob o ponto de vista pessoal quanto ao seu desempenho no curso superior. Empregando o método econométrico *Propensity Score Matching* (PSM), os autores constituíram dois grupos de estudantes denominados controle e tratamento, indicando, respectivamente, aqueles que não receberam benefícios e os assistidos com algum benefício dentre os concedidos pela IFES (auxílio permanência, alimentação, moradia e transporte), possibilitando medir as diferenças de desempenho entre os grupos. O coeficiente de rendimento global foi escolhido como indicador de desempenho acadêmico dos estudantes. Observou-se, como principal resultado, o impacto positivo da concessão dos auxílios sobre o desempenho dos estudantes, sendo que em algumas medições a diferença entre os grupos chegou a 16%.

As relações entre o PNAES, o desempenho acadêmico e a evasão escolar foram analisadas por [Cespedes et al. \(2021\)](#) dentro do contexto de uma das maiores IFES brasileiras. Os dois construtos foram comparados entre dois grupos de estudantes elaborados exclusivamente para a pesquisa, um com beneficiários de auxílios socioeconômicos e outro com não beneficiários, porém com o mesmo perfil de renda. A instituição define cinco perfis de vulnerabilidade socioeconômica associando renda e dados de análise social, de maneira que cada estudante seja classificado em um deles para subsidiar a justificativa de concessão dos auxílios. Os dados analisados referem-se ao período compreendido entre 2012 e 2016, totalizando 13.402 registros. A evasão considerou estudantes desligados formalmente da instituição, enquanto o indicador para medir o desempenho foi o coeficiente de rendimento semestral. Verificou-se menor percentual de evadidos entre o grupo de beneficiários, bem como maior desempenho acadêmico com dispersão mais equilibrada neste grupo. Ademais, os autores ressaltam que tal diferença no desempenho acadêmico ocorreu em todos os cinco perfis de vulnerabilidade, reforçando o papel do PNAES no cumprimento do objetivo de auxiliar na melhoria do desempenho acadêmico dos estudantes de IFES brasileiras.

[Oliveira, Deus e Campos \(2024\)](#) examinaram as contribuições do PNAES sobre a evasão e o desempenho acadêmico dos estudantes da Universidade Federal de Viçosa - Campus Rio Paranaíba. O período de estudo analisado foi de 2015 a 2019, englobando todos os dez cursos de graduação ofertados presencialmente. Os principais auxílios demandados pelos estudantes foram de alimentação e moradia, com crescente procura na sequência dos anos analisados. Quanto à manutenção do vínculo institucional, notou-se que somente 17% dos bolsistas evadiram de seus cursos. Sobre o desempenho acadêmico, mais de 62% dos estudantes atendidos pelo PNAES obtiveram coeficiente de rendimento superior a 60 pontos, mínimo para aprovação nas disciplinas. Os autores ressaltam, contudo, a necessidade de ampliação do alcance do apoio estudantil, de forma que um número maior de estudantes possa se beneficiar de condições favoráveis à sua permanência na instituição.

O estudo de [Silva, Freitas e Teixeira \(2024\)](#) propõe um instrumento para avaliar a política pública de assistência estudantil na Universidade Federal do Rio Grande (FURG) e o seu impacto em um dos indicadores abordados pelo PNAES, o desempenho acadêmico. A pesquisa utilizou a técnica *Propensity Score Matching* (PSM) para analisar os dados de estudantes matriculados entre 2014 e 2019, comparando o grupo daqueles beneficiados pelo programa (bolsistas) e dos não beneficiados. A variável de interesse analisada foi o coeficiente de rendimento acadêmico dos estudantes. Os resultados indicam que os beneficiários do PNAES superaram os não bolsistas em até 41,8% no coeficiente de rendimento, sugerindo que o programa contribui com o custeio dos bolsistas e com o incremento no seu desempenho.

O diferencial do estudo de [Barros \(2022\)](#) foi ter apresentado um recorte temporal abrangendo a situação do desempenho acadêmico de estudantes antes e durante a pandemia de COVID-19, possibilitando a comparação entre essas duas fases. A primeira fase, denominada pré-pandemia, envolveu três períodos letivos: 2018-2, 2019-1 e 2019-2. A mesma quantidade de períodos letivos constituíram a segunda fase: 2020-1, 2020-2 e 2021-1. As variáveis elencadas no conjunto de dados analisado envolveram aspectos demográficos, socioeconômicos e acadêmicos dos estudantes, sendo o coeficiente de rendimento acadêmico o indicador escolhido para representar o desempenho e a situação de matrícula igual a “cancelado” por semestres seguidos ou igual a “trancado” como indicador da evasão.

Uma das dificuldades relatadas no estudo de [Barros \(2022\)](#) foi justamente integrar estas variáveis em um único conjunto de dados. Assim como em outros estudos, os registros dos estudantes foram divididos em dois grupos, um representando os beneficiários de auxílios socioeconômicos e outro com os não assistidos. A verificação da influência do PNAES nas variáveis representantes dos outros construtos baseou-se na aplicação da

técnica conhecida por Diferenças em Diferenças (DD), cujo foco é prover avaliações do impacto de políticas públicas, em conjunto com validações estatísticas de significância. O estudo mostrou que o grupo de estudantes beneficiários possuía desempenho acadêmico superior ao obtido pelo grupo dos não assistidos, seja no recorte antes da pandemia ou durante a mesma. Resultado similar foi obtido no que tange à evasão, isto é, a quantidade de estudantes que trancaram ou cancelaram o curso, seja antes ou durante a pandemia, foi sempre menor entre aqueles do grupo de beneficiários, confirmando o PNAES como um dos fatores que auxiliaram na permanência estudantil.

3 METODOLOGIA

Este Capítulo apresenta a classificação metodológica da tese, bem como os procedimentos utilizados para alcance dos objetivos delineados no Capítulo 1.

3.1 Caracterização da pesquisa

Com o intuito de subsidiar a compreensão sobre a condução do desenvolvimento deste estudo, a modalidade da pesquisa (em suas diversas categorizações) foi definida conforme Gerhardt e Silveira (2009) e envolve a indicação das várias estratégias necessárias para o cumprimento dos objetivos propostos.

No que se refere à abordagem, foram utilizadas técnicas de pesquisa quantitativa para promoção de análises diversificadas sobre os dados coletados, envolvendo Estatística Aplicada e recursos de Aprendizado de Máquina, que habilitaram a geração de evidências quantificáveis alinhadas à proposta da tese. Ademais, em momentos específicos, algumas técnicas de pesquisa qualitativa, como grupo focal e entrevistas, contribuíram para promover uma visão mais subjetiva sobre determinados fatos. Quanto à natureza, trata-se de uma pesquisa aplicada que visa construir conhecimentos com utilidade prática em problemas relacionados aos três construtos (auxílio socioeconômico, evasão escolar e desempenho acadêmico) no âmbito do IFMG, em especial dos cursos presenciais de graduação. Quanto aos objetivos, tem-se uma pesquisa exploratória em função da demanda por levantamentos bibliográficos que subsidiaram a compreensão pelo pesquisador sobre as áreas e temas envolvidos. Trata-se, também, de uma pesquisa explicativa quanto aos objetivos, pois mostrou-se necessário analisar e entender as diversas variáveis advindas dos domínios dos construtos envolvidos, permitindo que fossem exploradas as características das relações entre elas e formuladas explicações coerentes sobre o impacto exercido por cada uma. Finalmente, no que tange aos procedimentos, o estudo consiste em uma pesquisa *Ex-Post Facto* a partir da demanda por investigações com dados reais de estudantes; pesquisa experimental, por simular variadas situações que subsidiaram a realização dos objetivos propostos; e estudo de caso, por ter apenas o IFMG como objeto de investigação.

[Moon e Blackman \(2014\)](#) propuseram um guia para que os pesquisadores compreendessem os princípios filosóficos das ciências sociais, auxiliando-os na caracterização de suas pesquisas sob uma abordagem multifacetada. De forma geral, existem três elementos fundamentais capazes de promover sentido aos projetos de pesquisa: **I) ontológico**, envolvendo tudo aquilo que existe no mundo e serve como objeto de investigação para pesquisadores; **II) epistemológico**, tratando questões sobre como o conhecimento é criado; **III) filosófico**, relacionada com as orientações que guiam as ações dos pesquisadores.

Quanto à ontologia, esta tese está vinculada com a concepção realista, caracterizada pela busca de “verdades” com o suporte de métodos científicos e com a elaboração de cenários que desagregam o todo no intuito de identificar propriedades que, por sua vez, habilitam a investigação de relações causais entre os sujeitos e objetos analisados. Quanto à epistemologia, tem-se um estudo no contexto da concepção objetivista, na qual os métodos são empregados para testar uma realidade em questão, envolvendo a coleta e análise de evidências com capacidade de justificar as afirmações da pesquisa e, portanto, transmitir confiabilidade quanto aos resultados obtidos.

Embora a tese adote uma concepção ontológica fundamentalmente realista, assumindo que há entidades e relações que existem independentemente da observação do pesquisador, optou-se pelo emprego da UFO como ontologia de fundamentação, reconhecendo seu papel enquanto arcabouço conceitual para representação formal de domínios do mundo real. Essa escolha não configura uma contradição epistemológica, dado que a UFO é suportada por um conceitualismo ontológico que, mesmo reconhecendo a mediação cognitiva nos processos de modelagem, é compatível com uma postura realista moderada ([GUIZZARDI, 2005](#)). Ademais, o uso de abordagens conceituais pode estar ancorado em compromissos ontológicos realistas, desde que se mantenha a distinção entre a realidade ontológica e sua representação ([GUIZZARDI et al., 2015](#); [ALMEIDA et al., 2023](#)). Dessa forma, a modelagem ontológica baseada na UFO não visa criar a realidade, mas representá-la com o maior grau possível de fidelidade conceitual, respeitando os limites impostos pela linguagem e pela cognição humana.

Quanto ao caráter filosófico desta tese, optou-se pela sua delimitação a partir da categorização proposta por Creswell (2014), justamente um dos autores que basearam o trabalho de [Moon e Blackman \(2014\)](#). Creswell (2014) afirma que a orientação filosófica das pesquisas científicas pode ser classificada com base em um agrupamento composto por quatro visões de mundo, ressaltado o caráter não excludente entre elas: **I) pós-positivista**, na qual existe a necessidade de se desafiar constantemente o que é tido como verdade absoluta, considerando que os resultados são determinados, possivelmente, em função de uma causa que precisa ser analisada no contexto de um experimento de observação e mensuração. A compreensão do mundo passa, portanto, pela verificação de relações causais em teorias com o suporte de dados e de considerações racionais, especialmente no contexto de pesquisas com abordagem quantitativa. **II) construtivista**, onde busca-se compreender qual é a complexidade de determinado problema por meio de uma visão menos específica e mais abrangente, construindo significados qualitativos para as questões pesquisadas a partir de discussões e interações promovidas entre indivíduos; **III) transformativa**, na qual verifica-se a intenção de abarcar assuntos marginalizados na sociedade e, em certa medida, no contexto acadêmico; **IV) pragmática**, onde existe a

preocupação em explorar a multiplicidade de aplicações e soluções para os problemas pesquisados, isto é, o que funciona para resolvê-los, em meio a uma busca por entendê-los melhor ao invés de dedicar demasiada concentração nos métodos, os quais podem ser combinados conforme as necessidades específicas da pesquisa.

Nesse caso, é possível classificar a pesquisa desta tese sob a ótica das concepções filosóficas denominadas pós-positivista e pragmática. A vinculação como pós-positivista baseou-se na demanda por investigações capazes de testar e avaliar confiavelmente a teoria que indica que determinado fator (a concessão de auxílio socioeconômico) é a causa influenciadora dos resultados (relativos ao desempenho acadêmico e à evasão escolar). A condução de investigações suportadas por dados reais se mostrou fundamental para a elaboração de considerações racionais promotoras de novos conhecimentos. A vinculação como pragmática adveio da necessidade do emprego de métodos mistos ajustados às demandas da pesquisa para que os problemas analisados tivessem variadas perspectivas de entendimento. É válido mencionar, ainda, o direcionamento da investigação para a compreensão dos problemas analisados e a ausência de aleatoriedade na combinação dos métodos quantitativos e qualitativos utilizados, cuja escolha seguiu a necessidade específica de determinados momentos da pesquisa quanto à melhoria no entendimento dos problemas investigados.

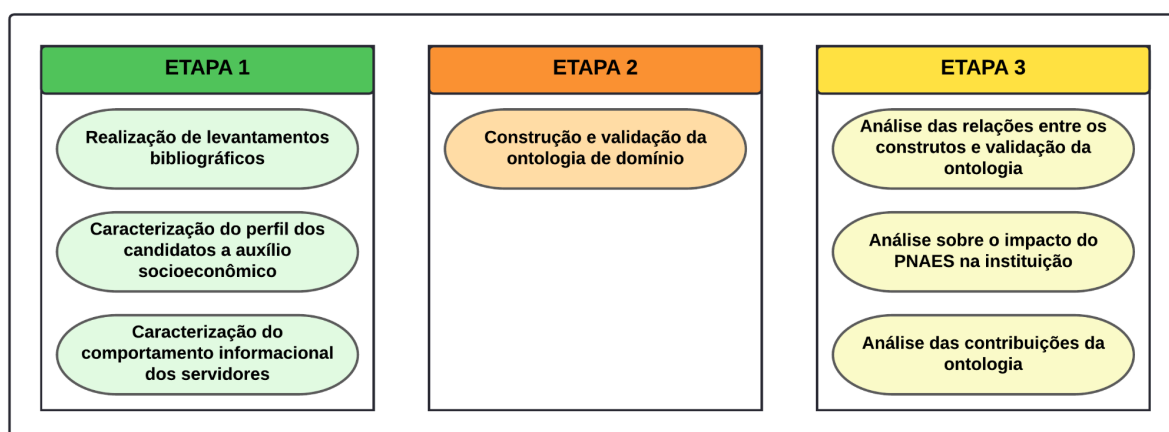
Outra caracterização da presente pesquisa é definida sob o ponto de vista da área de Computação e sua forte relação interdisciplinar com a Ciência da Informação ([PINHEIRO, 2008](#); [SARACEVIC, 1995](#)), na qual revela-se a importância das diversas técnicas, métodos e recursos computacionais aplicados na resolução de problemas informacionais como, por exemplo, a integração e processamento de grandes volumes de dados. Wazlawick (2009) entende que as pesquisas que permeiam a área de Computação (e outras correlatas) podem ser classificadas em cinco tipos: **I) Apresentação de um produto:** consiste na geração de novo produto a partir da realização de pesquisa exploratória, porém sem base científica anterior que permita a comparação dos resultados e o entendimento se o problema foi resolvido; **II) Apresentação de algo diferente:** envolve pesquisas, geralmente com enfoque qualitativo, onde busca-se introduzir uma forma diferente para resolver determinado problema, não necessariamente com a apresentação dos dados seguindo o devido rigor científico; **III) Apresentação de algo presumivelmente melhor:** abrange pesquisas que obtêm seus resultados a partir de comparações quantitativas entre uma nova abordagem para solucionar um problema e outras encontradas na literatura, sem foco na comprovação dos avanços eventualmente obtidos; **IV) Apresentação de algo reconhecidamente melhor:** são pesquisas com características similares àquelas do tipo III, mas que se diferenciam por “avançar o estado da arte” da área estudada, de maneira que os resultados encontrados permitam afirmar que a abordagem

proposta é mais vantajosa do que a formalizada por outros estudos científicos; **V) Apresentação de uma prova:** considera pesquisas que constroem uma teoria sobre determinado assunto e apresentam uma prova formal de que ela é correta. O estudo desenvolvido nesta tese pode ser tipificado, sob a ótica proposta por Wazlawick, como a apresentação de algo reconhecidamente melhor, pois os resultados contribuem para o estado da arte nas áreas de Educação e Ciências Sociais Aplicadas ao propor e divulgar evidências empíricas de uma solução diferenciada capaz de resolver um problema atualmente em aberto.

3.2 Procedimentos metodológicos

Para alcançar o resultado principal da tese, a pesquisa foi dividida em três etapas, cada uma contendo seu próprio conjunto de fases (Figura 14).

Figura 14: Etapas da pesquisa



Fonte: Elaborada pelo autor (2023)

3.2.1 Etapa 1

Esta etapa visava promover conhecimentos mais específicos acerca do contexto geral do estudo, os quais forneceram subsídios para a execução das etapas seguintes. Três produtos documentais foram gerados como resultados da etapa: o referencial teórico da pesquisa, um estudo sobre o perfil dos candidatos a auxílio socioeconômico e outro sobre o comportamento informacional dos servidores que participaram de processos seletivos para concessão de tais benefícios. A etapa é composta por três fases, expostas nas Subseções a seguir.

3.2.1.1 Realização de levantamentos bibliográficos

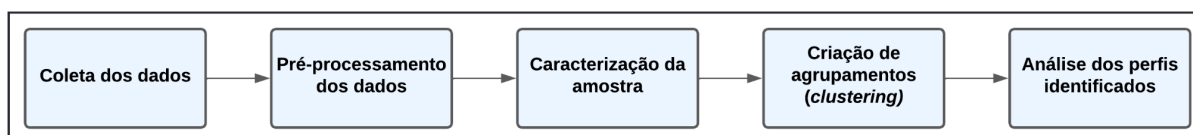
A primeira fase desta etapa consistiu na realização de levantamentos bibliográficos que permitiram maior aprofundamento teórico e embasamento conceitual sobre assuntos

vinculados com o tema central em questão, a saber: evasão escolar, desempenho acadêmico, assistência estudantil, gestão da informação, comportamento informacional, ontologias, mineração de dados, aprendizado de máquina e políticas públicas. De acordo com [Sousa, Oliveira e Alves \(2021\)](#), trata-se de uma estratégia que proporciona muitos benefícios para um estudo científico como, por exemplo, a ampliação do conhecimento sobre algum fenômeno e identificação de problemas ainda não respondidos acerca do mesmo. Com a análise crítica de publicações científicas nacionais e internacionais, buscadas por meio de palavras-chave relacionadas aos assuntos, o arcabouço teórico da pesquisa foi construído, dispondo as informações necessárias à compreensão dos vários conceitos que a suportam. As buscas se concentraram nas seguintes bases eletrônicas: *Scopus*, *Web of Science*, *ACM Digital Library*, *Scientific Electronic Library Online* (SciELO), *Educational Resources Information Center* (ERIC), Google Acadêmico e Biblioteca Digital Brasileira de Teses e Dissertações. Esta revisão da literatura serviu, ainda, como ponto de partida para a elaboração de uma seção descrevendo e relacionando os trabalhos correlatos ao aqui exposto, enfatizando suas similaridades e diferenças, tanto em termos de problemas quanto de soluções adotadas, para o que se pretendia construir nesta tese.

3.2.1.2 Caracterização do perfil dos candidatos a auxílio socioeconômico

Tendo como princípio que o entendimento sobre as características dos estudantes que participam de processos seletivos para concessão de auxílio socioeconômico é fundamental para compreender o público que demanda tal benefício na instituição, assim como a premissa de que tal caracterização fornece válidos subsídios para a construção da ontologia que suporta a integração dos dados pretendida neste estudo, a segunda fase da primeira etapa busca utilizar uma abordagem essencialmente quantitativa que permita, com o apoio de técnicas de Aprendizado de Máquina, criar agrupamentos dos candidatos com perfil semelhante. A Figura 15 exibe as atividades executadas nesta fase.

Figura 15: Atividades para caracterização do perfil dos candidatos a auxílios



Fonte: Elaborada pelo autor (2023)

A primeira atividade consistiu na coleta dos dados demográficos e socioeconômicos no sistema informatizado mantido pelo IFMG para o gerenciamento dos processos seletivos da Assistência Estudantil, onde são armazenados os dados dos candidatos e das suas participações nos editais publicados. Tal sistema foi concebido e desenvolvido pelo próprio

pesquisador desta tese, motivo pelo qual não será necessário o estudo mais aprofundado sobre as estruturas de armazenamento de dados utilizadas pela solução. Como o sistema mantém seus registros em um repositório do Sistema Gerenciador de Banco de Dados (SGBD) *MySQL* versão 5.7.36, foi criado outro repositório, nomeado *perfil_auxilios*, que recebeu apenas os dados necessários para o estudo em questão, recuperados por meio de consultas SQL (*Structured Query Language*). Todos os atributos possíveis dos candidatos foram coletados neste momento, sendo a filtragem daqueles mais relevantes para a caracterização definida na próxima atividade.

Após a disponibilização dos dados brutos em uma base específica, as tarefas de pré-processamento foram realizadas no escopo da segunda atividade, dentre as quais destacam-se a limpeza de registros eventualmente duplicados, verificação de dados faltantes, transformação de atributos em outros capazes de promover maior representatividade da informação do candidato e criação de novos atributos a partir de atributos já existentes ou baseados em enriquecimento promovido por fontes externas de dados. A realização de tais tarefas contou com o suporte dos softwares *phpMyAdmin* (versão 5.0.3) e *Visual Studio Code* (versão 1.59), em conjunto com *scripts Python* (versão 3.8). Finalmente, foi gerada uma tabela no repositório contendo o conjunto de dados final disponível para as análises propostas que, por sua vez, estão contempladas na terceira atividade. Para o seu cumprimento, promoveu-se a caracterização da amostra coletada com o suporte de recursos visuais para promover um entendimento mais claro sobre os dados, tais como gráficos e tabelas.

Ainda no contexto da atividade de caracterização do perfil dos candidatos e das solicitações de auxílio, compreendeu-se melhor o conjunto de dados com a identificação da importância de cada atributo para o deferimento das solicitações com a aplicação de uma técnica de *Feature Importance* conhecida como *Random Forest Feature Importance*, a qual consegue calcular e indicar numericamente a importância relativa de cada atributo na predição do resultado final. A importância de um atributo é calculada em cada árvore da floresta aleatória com base na sua capacidade de incrementar a pureza das folhas, de maneira que o resultado é diretamente proporcional a essa capacidade. Para reduzir eventuais variações, é calculada uma média entre todas as árvores para obtenção do indicador final de importância de cada atributo ([CHEN, 2024](#); [IRANZAD e LIU, 2024](#); [WANG et al., 2024](#)). Esta tarefa contou com o apoio do software *Visual Studio Code* (versão 1.59) e da biblioteca *scikit-learn* (versão 1.2.1).

A técnica de *Feature Importance* baseada em *Random Forest* foi escolhida por sua simplicidade de utilização e capacidade de avaliar a relevância dos atributos de forma robusta, isto é, pequenas mudanças nos dados não afetam muito o resultado da avaliação

dos atributos importantes, dada sua capacidade de analisar e combinar os resultados de diferentes árvores, tornando o resultado mais estável.

A quarta atividade envolveu o emprego de recursos de Aprendizado de Máquina, mais especificamente da categoria de aprendizado não supervisionado, para cumprir a tarefa de criar agrupamentos (*clustering*) dos estudantes em virtude do grau de semelhança entre suas características. Para efeitos de categorização dos procedimentos, a técnica escolhida é baseada em particionamento, o modelo criado executou o algoritmo *K-means* ([SAPUTRA, SAPUTRA e OSWARI, 2020](#)), os atributos foram normalizados na mesma ordem de grandeza e a quantidade ideal de agrupamentos inferida com o apoio do Método *Elbow*. A avaliação da performance do processo de agrupamento baseou-se no resultado da métrica *Silhouette Score*. O código-fonte elaborado foi executado no software *Visual Studio Code* (versão 1.59), contando com funcionalidades disponíveis nas bibliotecas *PyCaret* (versão 2.3.6), *pandas* (versão 1.5.3), *seaborn* (versão 0.12) e *scikit-learn* (versão 1.2.1).

O algoritmo *K-means* foi escolhido por sua simplicidade de implementação e eficiência computacional, especialmente em conjuntos de dados de tamanho moderado. Ele é amplamente utilizado em tarefas de agrupamento onde o número de *clusters* é estimável previamente, como no caso em questão.

Considerando o entendimento provido pela caracterização da amostra e o resultado da descoberta de agrupamentos, a quinta atividade promoveu análises quantitativas sobre os perfis identificados pela estratégia de Aprendizado de Máquina, permitindo compreender as características demográficas e socioeconômicas predominantes em cada agrupamento.

3.2.1.3 Caracterização do comportamento informacional dos servidores

De maneira complementar ao buscado com a fase anterior, na terceira fase da primeira etapa objetiva-se promover uma caracterização do comportamento informacional dos servidores da instituição pesquisada envolvidos com o processo de concessão auxílios estudantis, em especial daqueles que atuam na análise da documentação dos candidatos para comprovação da existência de vulnerabilidade social. Trata-se de uma seleção intencional da amostra, e esta depende do discernimento do pesquisador no sentido de que os elementos participantes sejam bons fornecedores de informação e possuam atributos que representem a população como um todo (MARCONI e LAKATOS, 2017).

A validade deste estudo reside no fato de que a atuação desses servidores é de extrema importância para todo o processo e altamente baseada na busca e uso de informações, posto que a decisão de atribuir (ou não) determinado benefício precisa ser devidamente lastreada no conjunto de dados não somente de um candidato de forma isolada, mas considerando diversas variáveis, como o perfil de vulnerabilidade dos outros

concorrentes, o recurso orçamentário disponível para distribuição dos auxílios, a quantidade de requerimentos, a suficiência do auxílio para o candidato, dentre outras.

O Modelo de Comportamento Informacional proposto com refinamentos por [Wilson \(1999\)](#) serviu como suporte teórico para a caracterização do comportamento informacional dos servidores da instituição. Os resultados da aplicação do Modelo estão disponíveis na literatura acadêmica em vários estudos referentes ao tema, fato que provê confiabilidade para sua escolha na tese. Conforme mencionado no Capítulo 2 (Referencial Teórico), trata-se de um Modelo que entende que as demandas de informação se originam de outras necessidades e surgem no âmbito de determinado contexto particular que, por sua vez, pode sofrer influência de variáveis capazes de intervir no processo.

A coleta dos dados que permitiram uma análise sobre o comportamento informacional dos servidores foi realizada com uma adaptação do questionário proposto por [Souza \(2020\)](#), acessível no Apêndice A. O artefato possui quinze questões (quatorze fechadas e uma aberta) e engloba as quatro categorias de contexto que compõem o construto Comportamento Informacional, conforme apontado por Choo (2006), [Wilson \(1999\)](#) e Davenport e Prusak (1998), a saber: I) **necessidade de informação**: procura levantar como os servidores identificam suas necessidades de informação; II) **busca da informação**: visa identificar quais são as práticas utilizadas pelos servidores para buscarem a informação, bem como as fontes envolvidas; III) **uso da informação**: objetiva mapear as práticas de uso da informação pelos servidores; IV) **compartilhamento da informação**: busca compreender como se dá o compartilhamento das informações a partir dos servidores. Em conjunto com as quatro categorias, o questionário aborda também a percepção dos envolvidos quanto aos atributos que qualificam a informação.

A aplicação do questionário ocorreu por meio eletrônico (*Google Forms*) para servidores de todos os *campi* da instituição que tenham participado de pelo menos um processo de seleção de bolsistas entre 2019 e 2022. Esse recorte no período, principalmente no que tange ao ano inicial, é necessário por duas razões: I) a padronização do processo e centralização da gestão na Reitoria ocorreram a partir de 2019, sendo que até então cada Campus conduzia seus processos de forma independente, inclusive com apoio de diferentes técnicas para coleta de dados (alguns utilizando planilhas eletrônicas e outros documentação em papel, para citar somente dois exemplos); II) a disponibilização de um sistema informatizado que auxilia na execução das atividades dos processos seletivos da assistência estudantil ocorreu também em 2019, permitindo que houvesse padronização não apenas dos procedimentos entre os *campi*, mas especialmente provendo a instituição com um ambiente tecnológico capaz de sistematizar, coletar e armazenar dados dos candidatos e dos processos para a concessão dos auxílios socioeconômicos.

Os dados preenchidos pelos respondentes foram sumarizados e tratados com auxílio do *software Microsoft Excel*, possibilitando a geração de gráficos e tabelas capazes de introduzir visualizações diferenciadas que contribuíram para a análise dos resultados e, conseqüentemente, para a promoção do conhecimento sobre as particularidades do comportamento informacional dos servidores.

Encerradas as atividades das três fases propostas na Etapa 1, têm-se os subsídios necessários para o início das rotinas previstas para a Etapa 2, detalhadas na próxima Subseção.

3.2.2 Etapa 2

Esta etapa visou construir uma estrutura tecnológica baseada em ontologia a fim de suportar a integração dos dados dos três construtos avaliados neste estudo. Tal estrutura é composta por uma ontologia de domínio que modela formalmente os principais conceitos e relações compreendidas nos domínios dos construtos, sendo a mesma validada com o apoio de uma ferramenta informacional adaptada para armazenar os dados conforme a organização do conhecimento proposta pela ontologia. A etapa constitui-se de apenas uma fase, exposta a seguir.

3.2.2.1 Construção e avaliação da ontologia de domínio

A construção da ontologia teve como suporte metodológico a abordagem elaborada por [FALBO, 2014](#), denominada *Systematic Approach for Building Ontologies* (SABiO). Todas as cinco fases da metodologia indicadas como “Processos de Desenvolvimento” estão contempladas, dada a necessidade de se elaborar uma ontologia de domínio composta pela ontologia de referência e pela ontologia operacional. A sequência de execução das fases privilegiou um ciclo iterativo e incremental, flexibilizando o cumprimento das atividades no sentido de que os seus resultados fossem revistos em diversos momentos do projeto. Ressalta-se o importante papel desempenhado pelos chamados “Processos de Apoio”, dentre os quais destacam-se a aquisição de conhecimentos (com forte apoio de especialistas de domínio) e a documentação de todos os produtos gerados ao longo da execução das atividades do projeto. Na sequência, são explicados os procedimentos metodológicos empregados em cada uma das cinco fases estruturadas pela SABiO.

Na primeira fase, nomeada como “Identificação de Propósito e Definição de Requisitos”, ocorreu a identificação, pelo pesquisador, do propósito da ontologia, bem como a eliciação dos seus principais requisitos e registro no formato de questões de competência, as quais são responsáveis por delimitar o escopo da ontologia, isto é, o que faz parte e o que não está contemplado no seu domínio. Ao final da fase, houve maior entendimento

sobre as razões que levaram à criação da ontologia e qual seria o seu escopo previsto para o final do ciclo de desenvolvimento.

A segunda fase (“Captura e Formalização da Ontologia”) visava elaborar a ontologia de referência e teve início com a análise, pelo pesquisador, da documentação interna do IFMG contendo detalhes sobre as estruturas de armazenamento de dados do sistema que controla rotinas envolvidas com processos gerenciais acadêmicos. Em seguida, foram disponibilizados pelo IFMG dois arquivos contendo os dados acadêmicos dos estudantes de cursos superiores de graduação presencial, um deles com dados caracterizadores e outro com os dados das disciplinas cursadas. Ambos os arquivos foram analisados pelo pesquisador com o intuito de complementar o entendimento sobre quais características poderiam ser modeladas na ontologia e, posteriormente, utilizadas como suporte às respostas das questões de competência.

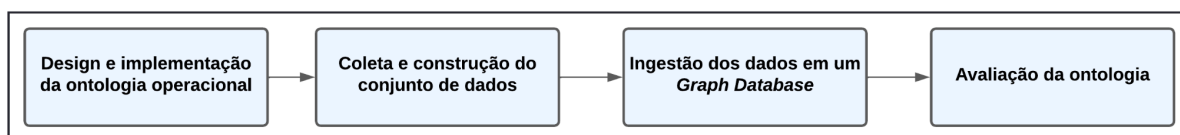
As atividades posteriores da segunda fase contaram com o apoio de duas experientes especialistas de domínio que atuam na instituição pesquisada, auxiliando por meio de entrevistas não estruturadas na compreensão dos conceitos e relações existentes nos domínios abordados na tese. Enquanto uma delas atua com o processo de concessão de auxílios socioeconômicos da Assistência Estudantil do IFMG, tendo sido gestora deste setor da instituição, a outra atua na área de ensino e tem experiência como Diretora de Ensino de um Campus. O pesquisador e as especialistas de domínio se reuniram virtualmente para elaborarem, com a análise dos resultados obtidos na primeira fase e nos estudos feitos no início da segunda fase, a especificação da ontologia mediante uma compreensão comum e definirem formalmente os conceitos envolvidos. Optou-se por escrever a definição desses conceitos com uso de linguagem natural a partir de bases teóricas que, por sua vez, validassem o entendimento que estava sendo registrado e contribuíssem para reduzir eventuais interpretações divergentes. A demanda por modularizar a ontologia também foi analisada nessa fase, com base nos critérios propostos pela SABiO: independência, coesão e tamanho. Possibilidades de reutilização de outros artefatos ontológicos fizeram parte de uma pesquisa conduzida para identificar fontes de recursos confiáveis que pudessem ser aproveitados nesta tese.

Para finalizar a segunda fase e permitir o estabelecimento de um consenso entre o pesquisador e as especialistas de domínio sobre a especificação da ontologia de referência, foi elaborado um modelo contendo os seus elementos e suas relações em conformidade com a linguagem OntoUML, um perfil que incorpora no Diagrama de Classes da UML alguns estereótipos específicos, fazendo com que a ontologia resultante esteja alinhada com a ontologia de fundamentação UFO. Em termos de ferramentas computacionais, o software *Visual Paradigm* versão 16.2 forneceu os recursos necessários para o cumprimento desta fase com a entrega da ontologia de referência.

As fases seguintes sugeridas pela SABiO (“Design” e “Implementação”), buscam produzir uma versão da ontologia de referência que possa ser operacionalizada, no sentido de leitura e uso, por máquinas. Para isso, é preciso que a especificação conceitual definida na ontologia de referência seja transformada no formato de uma especificação de projeto que considere, além das questões conceituais pertinentes, aspectos arquiteturais da plataforma informacional a ser utilizada para a implementação da ontologia operacional, bem como as restrições impostas por eventuais requisitos não funcionais tecnológicos. A metodologia SABiO acrescenta, em seguida, outra fase dedicada à realização de testes que visam avaliar a qualidade da ontologia operacional construída, bem como medir sua adequação enquanto base para a integração de dados de diferentes fontes. Esta avaliação, norteada pela capacidade da ontologia operacional implementada, visa encontrar respostas para as questões de competência definidas na primeira fase do ciclo.

Com o objetivo de simplificar o entendimento acerca de como a presente pesquisa conduziu as ações requeridas para cumprir o proposto pelas fases Design, Implementação e Testes da SABiO, um conjunto de atividades foi definido e pode ser visualizado na Figura 16. Na sequência é feito o detalhamento sobre a execução de cada atividade.

Figura 16: Atividades para construção do ambiente informacional integrado à ontologia



Fonte: Elaborada pelo autor (2023)

No escopo da primeira atividade (**Design e implementação da ontologia operacional**), houve a transformação da ontologia de referência em uma ontologia operacional no formato OWL usando gUFO e sintaxe *Turtle*, tarefa conduzida com o auxílio dos *softwares Visual Paradigm (e plugin para OntoUML⁵) e Protégé (versão 5.5.0)*.

A segunda atividade (**Coleta e construção do conjunto de dados**) tinha como requisito a elaboração da ontologia operacional que uniformiza o entendimento acerca dos conceitos e relações envolvendo os três construtos desta tese. A partir dela, tem-se portanto, a estruturação necessária para que os dados reais da instituição pesquisada fossem coletados e formatados de acordo com a padronização exigida pela ontologia. Atualmente, a instituição dispõe de dois diferentes sistemas informatizados que gerenciam os dados demandados pela tese: um da assistência estudantil, específico para controle dos processos de concessão de auxílios socioeconômicos; e outro para controle geral da sua área acadêmica, englobando dados dos cursos e dos estudantes. Tais sistemas não

⁵ <https://github.com/OntoUML/ontouml-vp-plugin>

possuem nenhum tipo de conexão direta ou interoperabilidade entre eles, ocasionando a demanda por tarefas para integração dos dados nessa fase. Os dados de ambos os sistemas estão armazenados em repositórios do SGBD *MySQL*, motivo pelo qual criou-se um repositório próprio com essa tecnologia em ambiente local para o presente estudo, responsável por receber os dados brutos selecionados nessas duas fontes. Visando enriquecer esses registros obtidos junto à instituição e armazenados localmente, houve também a integração com informações de repositórios externos, quais sejam a Plataforma Nilo Peçanha e a central de dados abertos do INEP⁶ (especificamente os conjuntos com dados acerca do ENADE e do Censo da Educação Superior).

Dada a amplitude dos cursos oferecidos pela instituição e a conseqüente variação nos perfis dos alunos em diferentes níveis de ensino, somado ao fato do PNAES ser um programa voltado para estudantes do ensino superior, optou-se por concentrar as análises nos dados de estudantes matriculados em cursos presenciais de graduação do IFMG cujo ingresso ocorreu entre o primeiro semestre letivo de 2018 e o segundo semestre letivo de 2022. Este recorte no período é necessário por três fatores: (I) a centralização dos dados da assistência estudantil na Reitoria aconteceu apenas a partir 2018, sendo que anteriormente os processos eram realizados de forma descentralizada por cada campus, o que inviabilizaria a coleta dos dados necessários para esta pesquisa; (II) cursos que possuem prazo mínimo de integralização em cinco anos tiveram pelo menos uma turma analisada por todo o ciclo; (III) abarcando as turmas ingressantes após o início de 2020, quando todo o contexto educacional foi alterado devido a pandemia de COVID-19, criou-se um recorte na pesquisa que permite a verificação de eventuais impactos deste momento na evasão escolar e no desempenho acadêmico dos estudantes.

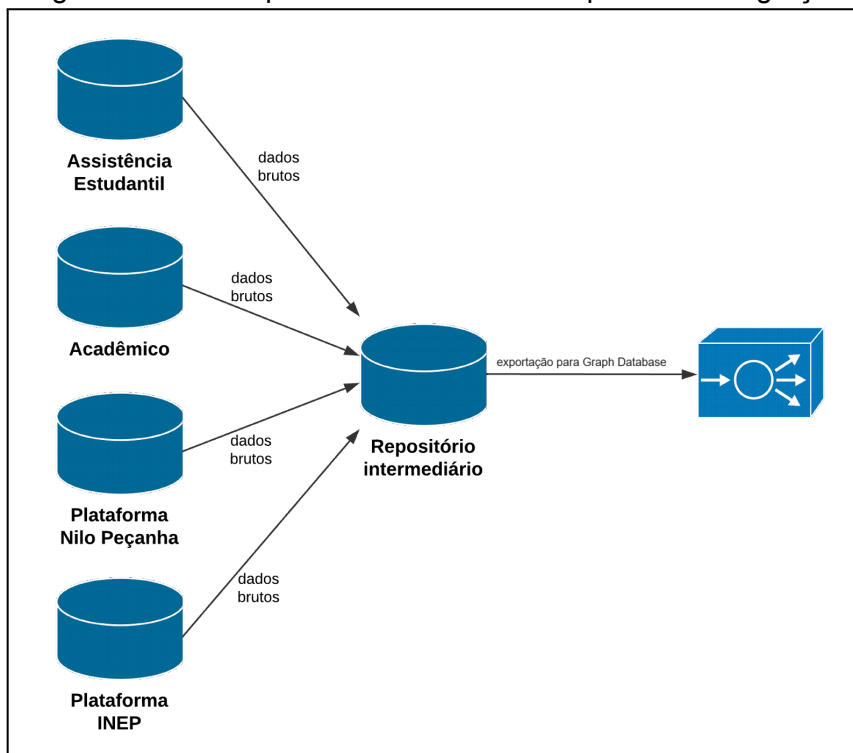
Quanto aos tipos de auxílios socioeconômicos, é considerada apenas a modalidade padronizada pela Reitoria do IFMG que atualmente é ofertada em todos os *campi*: bolsa permanência. Estão excluídos, portanto, aqueles concedidos apenas em alguns *campi* a partir de demandas específicas, como auxílio moradia e auxílio alimentação (no próprio Campus), bem como o auxílio inclusão digital, dado o fato de sua oferta ter sido restrita ao período inicial da pandemia de COVID-19.

Após a coleta dos dados e seu armazenamento em repositório próprio (intermediário), tarefas de pré-processamento foram executadas sobre os dados com o intuito de melhorar-lhes a qualidade e padronizá-los conforme requerido pela ontologia. Tais tarefas tiveram o suporte dos softwares *phpMyAdmin* (versão 5.0.3) e *Visual Studio Code* (versão 1.59), em conjunto com *scripts Python* (versão 3.8). Encerradas as manipulações, procedeu-se com o mapeamento entre os atributos do conjunto de dados e os conceitos da ontologia operacional elaborada na atividade anterior desta fase, habilitando a

⁶ <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/indicadores-educacionais>

transformação e exportação dos dados armazenados no repositório intermediário para o formato requerido pela plataforma de *Graph Database* (banco de dados em grafo). Na Figura 17 visualiza-se a importância da criação do repositório intermediário como ponto de integração dos dados advindos de diferentes fontes e como origem do conjunto de dados a ser manuseado na próxima atividade.

Figura 17: Fontes para coleta dos dados e ponto de integração



Fonte: Elaborada pelo autor (2023)

Na terceira atividade (**Ingestão dos dados em um *Graph Database***), os dados foram extraídos do repositório intermediário com a execução de consultas SQL e salvos em arquivos CSV estruturados conforme os elementos integrantes da ontologia operacional implementada em OWL. Optou-se pela escolha da plataforma *Neo4j*⁷ como ambiente para o gerenciamento (persistência e manipulação) do banco de dados no formato de grafo, cuja versão para uso livre e sem custos disponibiliza todas as funcionalidades necessárias para o desenvolvimento das tarefas desta tese, além de proporcionar alta performance na execução de ações (como leitura e escrita de dados) e ter baixa curva de aprendizado (SCIFO, 2023). A importação dos dados armazenados nos arquivos CSV para o repositório em grafo demandou a execução de comandos específicos da linguagem *Cypher*, a qual é utilizada no software *Neo4j Browser* versão 5.12, gerando como resultado final a criação de todas as instâncias e suas relações no projeto do *Neo4j*.

⁷ <https://neo4j.com/>

Embora a ontologia de referência tenha sido construída com base em OntoUML e, na sequência, convertida em OWL, o que naturalmente permitiria sua operacionalização por meio de tecnologias como RDF e SPARQL, optou-se, nesta tese, pela utilização de um modelo alternativo de representação de dados: o grafo de propriedades rotuladas (*label property graph*), operacionalizado com a ferramenta *Neo4j*. Esta decisão foi baseada em dois fatores: performance e simplicidade na formulação de consultas ([KOTIRANTA, JUNKKARI e NUMMENMAA, 2022](#)). Portanto, a escolha pelo *Neo4j* não representa rejeição às abordagens tradicionais baseadas em OWL/RDF/SPARQL, mas tão somente uma decisão estratégica orientada à eficiência e usabilidade, considerando as características dos dados e os objetivos analíticos desta investigação.

Para a execução da última atividade (**Avaliação da ontologia**), a análise da qualidade da ontologia construída e sua aptidão para integrar os dados dos três construtos consistiu em avaliar a sua capacidade para responder as questões de competência elaboradas no início do projeto. Cada questão foi implementada como uma consulta no contexto de um caso de teste, envolvendo a pergunta em si, os dados de entrada e a saída esperada. Objetivando validar estatisticamente se eram significativas (p abaixo de 0,05) as diferenças quantitativas encontradas entre os resultados das questões de competência que envolviam comparação entre o desempenho acadêmico de bolsistas e não bolsistas, aplicou-se o Teste T de *Student* (ou apenas “Teste T”) nos dados envolvidos em cada uma delas, pois essa abordagem permite a comparação das diferenças entre as médias de duas amostras (ASSIS, SOUSA e LINHARES, 2020). O *software Microsoft Excel* serviu como ferramenta para a execução de tais testes. Houve, também, a construção de uma matriz de rastreabilidade para verificar aqueles elementos da ontologia (conceitos, relacionamentos ou propriedades) que participaram da resolução de cada questão de competência, conforme proposto em [Silveira et al. \(2021\)](#), permitindo visualizar o alinhamento entre as questões e a ontologia, bem como delimitar, em alto nível, a cobertura do domínio pelas questões.

Quanto aos processos de apoio definidos na SABIÓ, é válido mencionar a utilização do *GitHub* como ferramenta para gerenciamento das versões dos artefatos produzidos pelas atividades deste estudo, sejam eles documentais, tecnológicos ou de dados. No que tange ao processo de aquisição de conhecimentos, destaca-se o alinhamento com o suporte cognitivo das especialistas de domínio participantes da pesquisa e o apoio de documentos internos da instituição pesquisada.

Encerrada a segunda etapa, estavam cumpridos os requisitos necessários ao início das atividades que compõem a terceira etapa (detalhada na próxima Subseção).

3.2.3 Etapa 3

Nesta etapa foram conduzidas três fases de análises orientadas a dados para promover o alcance dos objetivos gerais da tese, tanto o socioeconômico quanto o técnico. A realização de tais análises somente foi viável dada a prévia disponibilização de um ambiente informacional hábil para integrar os dados dos construtos com a devida qualidade em termos de organização proporcionada por uma ontologia. Os detalhes de cada fase são expostos nas próximas Sub-Seções.

3.2.3.1 Análise das relações entre os construtos e avaliação da ontologia

O escopo da primeira atividade desta fase abrangeu o desenvolvimento de um protótipo de aplicação computacional no formato de *dashboard* gerencial (ECKERSON, 2010) com recursos de visualização de informações envolvendo os construtos abordados na tese. Este *dashboard* apresenta as informações em um ambiente dinâmico, possibilitando aos usuários a montagem de suas próprias visões dos dados e a geração de novos conhecimentos. Quanto aos aspectos tecnológicos do protótipo, sua construção ocorreu no software *Microsoft Power BI Desktop*⁸ (versão 2.105) em conjunto com a plataforma do *Neo4j*, a qual fornece a estrutura de armazenamento dos dados. A integração deste software com a plataforma é possível graças a um recurso chamado *Neo4j Connector for BI*, habilitado para fornecer acesso direto aos dados armazenados em grafos por ferramentas de *Business Intelligence* (BI).

Tanto o pesquisador quanto as duas especialistas de domínio do IFMG (as mesmas que contribuíram com o processo de elaboração da ontologia) construíram, conjuntamente, análises quali-quantitativas sobre a relação entre os construtos a partir do acesso às informações disponibilizadas pelo *dashboard* gerencial e pelos resultados das questões de competência, as quais foram compartilhadas e discutidas em uma sessão de grupo focal realizada virtualmente. No entendimento de [Oliveira e Vasconcellos \(2021\)](#), a técnica de grupo focal envolve um trabalho conjunto com o propósito de obter novas informações a partir da interação entre os participantes, complementando eventuais resultados prévios obtidos por determinada pesquisa. Para [Boccatto e Ferreira \(2014\)](#), um diferencial importante trazido pelo grupo focal reside no fato de que as informações produzidas em conjunto possuem maior profundidade, diversidade e riqueza de detalhes em comparação com outras técnicas de coleta e análise de dados aplicadas individualmente. A aderência desta técnica à pesquisa fica latente quando considerada a necessidade de construção de novos conhecimentos a partir de reflexões de especialistas sobre os dados quantitativos ofertados pelo *dashboard* gerencial e pelas respostas às questões de competência.

⁸ <https://powerbi.microsoft.com/pt-br/>

3.2.3.2 Análise sobre o impacto do PNAES na instituição

O propósito dessa fase era analisar e mensurar o impacto no IFMG do PNAES enquanto Política Pública, utilizando como método uma abordagem conhecida como pareamento baseado no escore de propensão (*Propensity Score Matching - PSM*). Trata-se de uma técnica desenvolvida por [Rosenbaum e Rubin \(1983\)](#) que vem sendo utilizada para avaliar a causalidade de políticas e programas governamentais na sociedade, como observa-se em diversos estudos, tais como [Caliendo e Kopeinig \(2008\)](#), [França e Gonçalves \(2010\)](#), [Cavalcanti \(2015\)](#), [Salvini et al. \(2019\)](#) e [Gomes Neto, Santos e Prates \(2020\)](#). Embora na Subseção anterior tenham sido elaboradas análises sobre as relações entre os construtos no âmbito do IFMG, a aplicação do PSM permitiu o emprego de um método estatístico formal para auxiliar na investigação do impacto do PNAES na instituição.

A avaliação da política pública referida nessa tese (PNAES) é classificada como *ex-post* de impacto, pois sua realização ocorreu após o início da implementação das ações da política e visa compreender se ela produziu impacto nos problemas dos indivíduos envolvidos ([DRAIBE, 2001](#)).

De acordo com [Ramos \(2009\)](#), o PSM é utilizado em avaliações que objetivam averiguar possíveis efeitos causais de um tratamento específico sobre algum resultado. Inicialmente, ocorre a identificação de elementos em uma amostra de não participantes do tratamento, alocando-os em um grupo de controle, com a ressalva que as características observáveis dos elementos desse grupo sejam similares àquelas verificadas no grupo de tratamento. Cria-se, portanto, um grupo contrafactual ao grupo no qual foi aplicado o tratamento, habilitando a realização de comparações mais realistas e equilibradas.

[Resende e Oliveira \(2008\)](#) complementam que a geração de um grupo de controle com grande similaridade ao grupo de tratamento torna mais crível, a nível estatístico, a comparação entre as médias dos dois grupos e a estimação dos impactos do tratamento aplicado. No caso da pesquisa conduzida nesta tese, as características observáveis, nomeadas pelo método como variáveis independentes, poderiam ser a idade ao ingressar no curso, renda *per capita*, cor e forma de ingresso, enquanto a situação de um estudante como bolsista nos programas de assistência estudantil do IFMG é a variável de tratamento. O desempenho acadêmico (indicado pelo coeficiente de rendimento) e a situação de matrícula igual a evadido seriam as variáveis dependentes.

[Rosenbaum e Rubin \(1983\)](#) explicam que a implementação do PSM inicia-se com o cálculo de uma regressão logarítmica para estimar o escore de propensão, uma variável de controle que indica a probabilidade de determinado indivíduo receber o tratamento em questão, contanto que algumas variáveis observáveis sejam consideradas. Estas variáveis representam as características de determinada amostra e, supostamente, contribuem para o motivo daquele indivíduo ser tratado ou não. [Resende e Oliveira \(2008\)](#) ressaltam que é

necessário haver um equilíbrio em relação à quantidade de variáveis que caracterizam a amostra, pois quando elas são muitas, pode ser mais complexo para o método identificar indivíduos no grupo de controle com, literalmente, as mesmas características de indivíduos que compõem o grupo de tratamento.

Realizado o cálculo para estimar o escore de propensão, é preciso examinar se existe balanceamento entre os grupos de controle e tratamento. Para isso, foi executado o Teste T de Student e encontrada a diferença das médias em cada variável independente, tanto antes como depois do pareamento. A geração de gráficos de densidade permitiu visualizar se os escores de propensão estimados para os grupos estavam equilibradamente distribuídos, então comprovando a existência de uma região de suporte comum entre eles.

[Reis et al. \(2015\)](#) afirmam que, na prática, é difícil que o escore de propensão seja exatamente o mesmo para dois indivíduos comparados, motivo pelo qual é necessário estimar o *Average Treatment Effect on the Treated* (ATT - Efeito Médio de Tratamento sobre o Tratado) com o auxílio de um sistema de pesos. O método do vizinho mais próximo pode funcionar como esse sistema auxiliar, buscando no grupo de tratamento um indivíduo com igual (ou mais próximo) escore de propensão para cada caso do grupo de controle, tendo ao final dois grupos com indivíduos pareados. O cálculo do ATT utiliza, por padrão, os cinco vizinhos mais próximos a cada indivíduo analisado.

Para encerrar o processo de utilização do PSM, ocorre a estimação do impacto do tratamento sobre as variáveis dependentes com duas medidas: o ATT e a diferença de médias dos resultados dos grupos com verificação de significância estatística pelo Teste T. A execução de todos os testes mencionados nesta fase contou com o suporte do software *RStudio* versão 2023.06.2, além do *Microsoft Excel* para preparação dos dados.

No próximo Capítulo são descritos e analisados os resultados obtidos ao final da execução de todas as atividades relacionadas nos procedimentos metodológicos.

4 RESULTADOS E DISCUSSÕES

Este Capítulo apresenta em detalhes o desenvolvimento das atividades especificadas na Metodologia, bem como discute os resultados obtidos. A Seção 4.1 promove uma caracterização do perfil dos candidatos a auxílio socioeconômico do IFMG, enquanto, na Seção 4.2, são analisadas as características do comportamento informacional dos servidores da instituição que atuam no setor de assistência estudantil. A Seção 4.3 apresenta o detalhamento das atividades necessárias para a elaboração da ontologia que suportou a integração dos dados utilizados nesta pesquisa, incluindo a sua verificação. Na Seção 4.4, as relações entre os construtos foram discutidas com base nos dados disponibilizados por um *dashboard* gerencial, seguidas da validação da ontologia elaborada por especialistas de domínio. Por fim, na Seção 4.5, foi promovida uma análise formal sobre o impacto do PNAES na instituição.

4.1 Caracterização do perfil dos candidatos em processos seletivos da assistência estudantil

A primeira atividade desta fase envolveu a coleta dos dados demográficos e socioeconômicos dos candidatos vinculados a cursos de graduação que participaram de processos seletivos da assistência estudantil do IFMG ofertados desde o início de 2019 até o final de 2022. Por meio de uma consulta SQL, os dados brutos foram obtidos e diretamente armazenados em um repositório no SGBD *MySQL* identificado como *perfil_auxilios*. É importante ressaltar que houve uma solicitação formal por parte do pesquisador junto ao IFMG para acesso aos dados, a qual foi aprovada com a condição de anonimização de todos os registros obtidos.

Em seguida, passou-se à atividade de pré-processamento dos dados, cujas manipulações foram realizadas diretamente no repositório do *MySQL*. Não foram encontrados registros duplicados que demandassem exclusão, nem reconhecida a necessidade de integração com outras bases internas de dados. Dado o fato de o sistema informatizado utilizado pela instituição requerer o preenchimento obrigatório de todos os campos pelos candidatos, sendo que tal ação ocorre por meio de seleção de valores previamente apresentados na maioria dos campos, não foi identificado nenhum dado faltante com demanda para atribuição de valor. Analisando a estrutura do conjunto de dados, identificou-se a possibilidade de transformar alguns atributos originais em outros com valores mais relevantes para o estudo: *idade_na_inscricao* (obtido a partir de um cálculo envolvendo a data de nascimento do candidato e a data da inscrição no edital) e *nascido_em_outro_estado* (obtido com a verificação se o estado de nascimento do candidato é igual a Minas Gerais). Os atributos que originaram estas transformações foram

excluídos da base de dados ao final das manipulações. Três novos atributos foram incluídos no conjunto de dados: *area_curso* (indicando a área do curso de acordo com a [Tabela de Áreas de Conhecimento/Avaliação da CAPES](#)); *modalidade_curso* (representando a modalidade do curso vinculado com o candidato, podendo ser Bacharelado, Licenciatura ou Tecnólogo); e *turno_curso* (contendo qual o turno de oferta do curso, dentre diurno, noturno ou integral).

Ao término das atividades de preparação, o conjunto de dados ficou com 7.366 registros, compostos de 32 atributos (27 originais e 5 criados), os quais são nomeados e detalhados na Tabela 5. Foi realizada a exportação do conjunto de dados no formato CSV (*Comma-Separated Values*) para utilização na atividade seguinte.

Tabela 5 - Atributos dos candidatos a auxílios socioeconômicos

Nº	Atributo	Tipo	Criado?	Descrição/valores possíveis
1	campus	Nominal	Não	Nome do Campus de vínculo do estudante.
2	curso	Nominal	Não	Nome do curso do estudante.
3	area_curso	Categórico	Sim	Área do curso na CAPES. 0 (Ciências Agrárias) 1 (Ciências Biológicas) 2 (Ciências Exatas e da Terra) 3 (Ciências Humanas) 4 (Ciências Sociais Aplicadas) 5 (Engenharias) 6 (Linguística, Letras e Artes) 7 (Ciências da Saúde)
4	modalidade_curso	Categórico	Sim	Modalidade do curso do estudante. 0 (Bacharelado) 1 (Licenciatura) 2 (Tecnólogo)
5	turno_curso	Categórico	Sim	Turno de oferta do curso do estudante. 0 (Diurno - Matutino ou Vespertino) 1 (Noturno) 2 (Integral)
6	ano_edital	Numérico	Não	Ano de publicação do edital.
7	identidade_genero	Categórico	Não	Identidade de gênero do estudante. 0 (Feminino) 1 (Masculino) 2 (Outra)
8	estado_civil	Categórico	Não	Estado civil do estudante. 0 (Casado) 1 (Solteiro) 2 (Outro)
9	cor	Categórico	Não	Cor do estudante. 0 (Branco) 1 (Preto) 2 (Pardo) 3 (Amarelo/Indígena)
10	deficiencia	Categórico	Não	Indica se o estudante possui algum tipo de deficiência. 0 (Não) 1 (Sim)

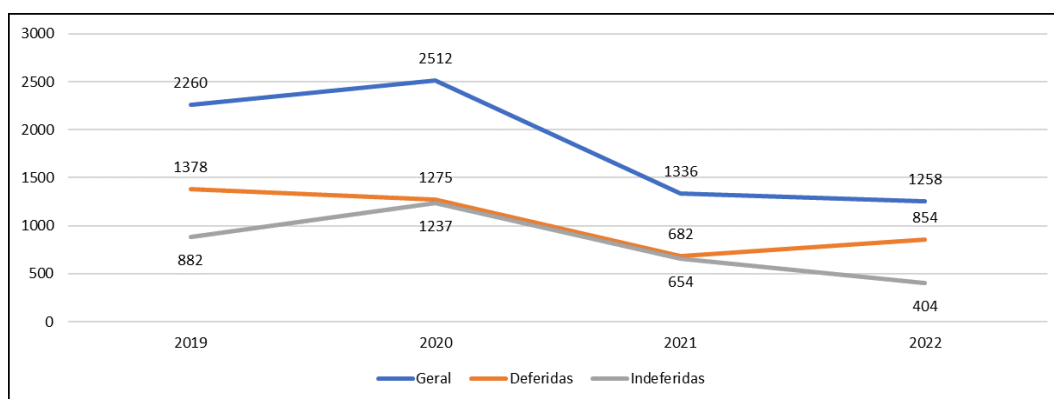
11	instituicao_ensino_fund	Catagórico	Não	Tipo da instituição onde o estudante cursou o Ensino Fundamental. 0 (Pública) 1 (Particular) 2 (Ambas)
12	instituicao_ensino_med	Catagórico	Não	Tipo da instituição onde o estudante cursou o Ensino Médio. 0 (Pública) 1 (Particular) 2 (Ambas)
13	forma_ingresso	Catagórico	Não	Forma de ingresso do estudante no curso. 0 (Ações afirmativas / cotas) 1 (Ampla concorrência)
14	escolaridade	Catagórico	Não	Escolaridade do estudante. 0 (Médio completo) 1 (Superior incompleto) 2 (Superior completo)
15	situacao_trabalho	Catagórico	Não	Situação de trabalho do estudante. 0 (Não trabalha) 1 (Trabalha)
16	nascido_em_outro_estado	Catagórico	Sim	Indica se o estudante nasceu em algum Estado fora de Minas Gerais. 0 (Não) 1 (Sim)
17	orgf_problemas_saude	Catagórico	Não	Indica se algum membro familiar do estudante possui problemas de saúde. 0 (Não) 1 (Sim)
18	situacao_moradia	Catagórico	Não	Situação de moradia do estudante. 0 (Moro com familiares) 1 (Moro sozinho) 2 (República/Moradia Estudantil)
19	situacao_imovel	Catagórico	Não	Situação do imóvel onde o estudante reside. 0 (Próprio) 1 (Alugado) 2 (Outra)
20	meio_transporte_para_campus	Catagórico	Não	Meio de transporte utilizado pelo estudante para deslocamento ao Campus. 0 (Próprio) 1 (Público) 2 (Não utilizo)
21	idade_na_inscricao	Numérico	Sim	Idade no estudante no momento da inscrição no processo seletivo.
22	possui_veiculo	Catagórico	Não	Indica se o estudante possui veículo. 0 (Não) 1 (Sim)
23	possui_imoveis_alem_moradia	Catagórico	Não	Indica se o estudante (ou o grupo familiar) possui imóveis além daquele utilizado como moradia. 0 (Não) 1 (Sim)
24	recebeu_auxilio_ano_anterior	Catagórico	Não	Indica se o estudante recebeu auxílio socioeconômico da AE no ano anterior à solicitação. 0 (Não) 1 (Sim)
25	recebe_bolsa_pesquisa_extensao	Catagórico	Não	Indica se o estudante recebe alguma bolsa como membro de projeto de pesquisa ou extensão. 0 (Não)

				1 (Sim)
26	recebe_bolsa_monitoria_tutoria	Catagórico	Não	Indica se o estudante recebe alguma bolsa como monitor ou tutor de disciplina. 0 (Não) 1 (Sim)
27	num_pessoas_gf	Numérico	Não	Número de pessoas que compõem o grupo familiar do estudante (inclusive com ele).
28	responsavel_gf	Catagórico	Não	Indica se o estudante é o responsável pelo seu grupo familiar. 0 (Não) 1 (Sim)
29	renda_per_capita	Numérico	Não	Renda <i>per capita</i> do estudante.
30	ivs	Numérico	Não	Índice de Vulnerabilidade Social (calculado conforme regras da Política de AE do IFMG).
31	tipo_bolsa	Catagórico	Não	Tipo da bolsa atribuída ao estudante ao final do processo de análise. Bolsa permanência 1 Bolsa permanência 2 Bolsa permanência 3 Bolsa permanência 4
32	situacao_final	Catagórico	Não	Indica se o estudante foi contemplado com algum auxílio ao final do processo seletivo. 0 (Não) 1 (Sim)

Fonte: Elaborada pelo autor (2023)

A terceira atividade desta fase buscou caracterizar a amostra resultante da atividade anterior, promovendo visualizações que permitam compreender os aspectos sociais e demográficos dos candidatos. Quanto às solicitações de auxílio (bolsa permanência), observa-se na Figura 18 que, embora tenha havido crescimento de 2019 para 2020, a quantidade de requisições nos anos seguintes foi notadamente menor. É possível perceber, ainda, que em 2019 e 2022 a diferença de solicitações deferidas para aquelas indeferidas foi perceptível, enquanto nos dois anos seguintes os números são mais próximos.

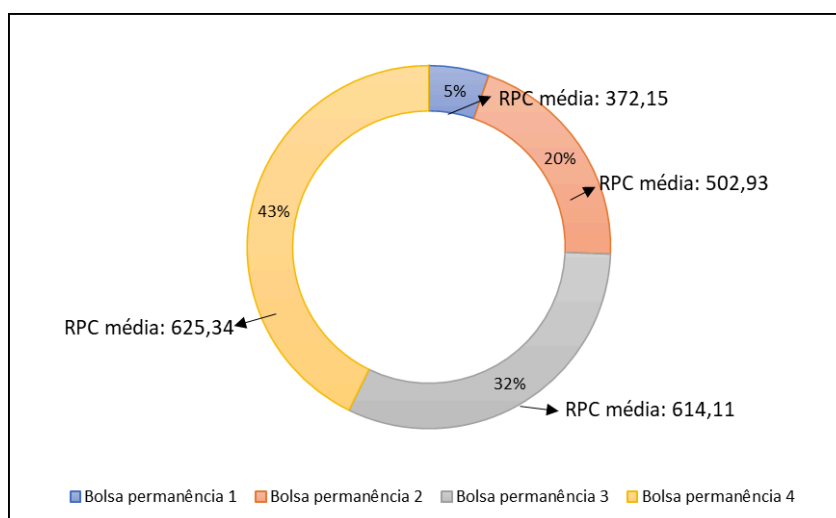
Figura 18: Solicitações de auxílios por ano do edital



Fonte: Elaborada pelo autor (2023)

Analisando a Figura 19, verifica-se que a maior parte das bolsas distribuídas aos candidatos contemplados se concentra nos auxílios BP4 (43%) e BP3 (32%), os quais pagam os menores valores (R\$150,00 e R\$200,00, respectivamente). O auxílio BP1, cujo valor da bolsa é de R\$400,00, é distribuído para apenas 5% dos contemplados. Entretanto, há de se considerar o fato de que a renda *per capita* média deste grupo é a menor, indicando a pertinência de se atribuir um auxílio de valor mais alto para aqueles candidatos com renda mais baixa. É possível visualizar certo padrão na distribuição do auxílio em relação à renda: valores menores para candidatos com rendas maiores e valores maiores para candidatos com menores rendas.

Figura 19: Distribuição geral das bolsas



Fonte: Elaborada pelo autor (2023)

Alguns detalhes das solicitações de auxílio por Campus são apresentados na Tabela 6. Estão expostas a quantidade de solicitações, a quantidade e o percentual de solicitações deferidas e indeferidas, a bolsa mais distribuída e a renda *per capita* média dos candidatos que tiveram solicitações deferidas e indeferidas. Os *campi* Bambuí, São João Evangelista e Formiga são os que tiveram maior quantidade de solicitações, mas apenas os dois primeiros ultrapassaram a marca de 1.000 pedidos de auxílio. Este indicador varia muito entre os *campi*, dadas as particularidades que os diferenciam, tais como cursos ofertados e ano de início de funcionamento. Embora 57% de todas as solicitações tenham sido deferidas, nos *campi* Arcos e Ibirité a quantidade de indeferimentos foi maior. A bolsa BP3, cujo valor pago é de R\$335,00, foi a mais distribuída entre os *campi*, totalizando oito unidades. Chama a atenção o percentual de distribuição da bolsa BP4 nos *campi* Ponte Nova e São João Evangelista, alcançando 97% e 93%, respectivamente. Quando a renda *per capita* dos candidatos que obtiveram deferimento é comparada com a daqueles que receberam

indeferimento de suas solicitações, observa-se que a instituição segue o preconizado pelo PNAES, pois a renda média dentre os deferidos é menor em todos os *campi*.

Tabela 6: Solicitações de auxílios por Campus

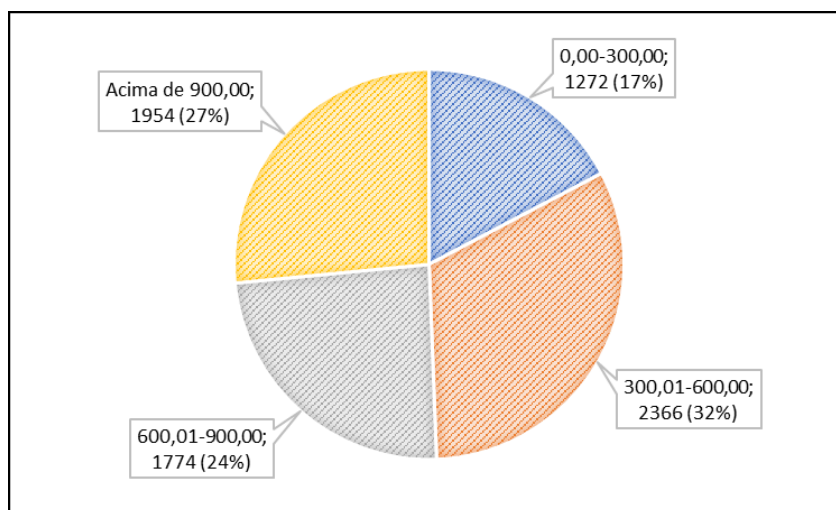
Campus	Bolsa mais distribuída	RPC média deferidas	RPC média indeferidas
Arcos	BP 2 (44%)	560,80	876,74
Bambuí	BP 3 (47%)	607,92	845,75
Betim	BP 2 (36%)	612,88	967,37
Congonhas	BP 3 (29%)	662,74	886,60
Formiga	BP 3 (47%)	566,17	873,28
Governador Valadares	BP 3 (42%)	652,68	872,86
Ibirité	BP 4 (47%)	627,38	1.102,83
Ipatinga	BP 2 (27%)	659,99	1.030,41
Itabirito	BP 3 (28%)	833,38	1.258,42
Ouro Branco	BP 2 (30%)	633,56	963,67
Ouro Preto	BP 4 (76%)	600,20	779,74
Piumhi	BP 4 (39%)	629,50	937,77
Ponte Nova	BP 4 (97%)	560,36	805,52
Ribeirão das Neves	BP 4 (48%)	578,75	1.106,16
Sabará	BP 3 (40%)	580,06	1.101,12
Santa Luzia	BP 3 (39%)	691,59	889,72
São João Evangelista	BP 4 (93%)	382,15	632,59
Totais	BP3 (freq: 8)	614,12	937,09

Fonte: Elaborada pelo autor (2023)

Apenas para efeitos de comparação, a renda *per capita* dos candidatos foi classificada em quatro categorias: de zero a trezentos reais, acima de trezentos reais até seiscentos reais, acima de seiscentos reais até novecentos reais, acima de novecentos reais. Verificando a Figura 20, embora exista certo equilíbrio entre o quantitativo das categorias, a maioria dos candidatos (32%) dispõe de renda *per capita* entre trezentos e seiscentos reais, enquanto o menor percentual (17%) tem renda de até trezentos reais. O PNAES estipulava, até 2024, como critério básico para priorização da concessão de auxílios socioeconômicos, renda *per capita* do candidato de até um salário mínimo e meio. Considerando que em 2022 este valor correspondia a R\$1.818,00 (um mil, oitocentos e

dezoito reais), 7.160 solicitações estavam dentro deste limite, ao passo que 206 solicitações eram de candidatos com renda mais elevada, das quais 14 foram deferidas.

Figura 20: Solicitações por renda *per capita*



Fonte: Elaborada pelo autor (2023)

Quando o número de solicitações de bolsas e a quantidade de deferimentos e indeferimentos são analisados por curso, os valores apresentam grande variação entre aqueles ofertados na instituição. A Tabela 7 exibe os cinco cursos com mais solicitações em todo o IFMG. O curso de Administração do Campus Ribeirão das Neves é o que recebeu o maior número de solicitações no período da pesquisa, com taxa de deferimento superior a 54%. São João Evangelista é o segundo Campus com maior quantidade de solicitações (1.140), sendo que pouco mais da metade está concentrada nos dois cursos relacionados. Destaca-se, também, o fato de os quatro primeiros cursos serem ofertados na modalidade Bacharelado.

Tabela 7: Cinco cursos com mais solicitações de bolsas

Campus	Curso	Modalidade	Deferidas	Indeferidas	Total
Ribeirão das Neves	Administração	Bacharelado	166 (54,2%)	140 (45,8%)	306
São João Evangelista	Agronomia	Bacharelado	184 (60,9%)	118 (39,1%)	302
Governador Valadares	Engenharia de Produção	Bacharelado	160 (55,6%)	128 (44,4%)	288
São João Evangelista	Engenharia Florestal	Bacharelado	154 (55,2%)	125 (44,8%)	279
Ouro Branco	Pedagogia	Licenciatura	140 (53,2%)	123 (46,8%)	263

Fonte: Elaborada pelo autor (2023)

Na Tabela 8 estão relacionados os cinco cursos com o menor quantitativo de solicitações de bolsas, entretanto é válido ressaltar que o curso de Educação Física do Campus Bambuí começou sua oferta em 2022, último ano compreendido no escopo desta pesquisa, enquanto o curso de Licenciatura em Computação do Campus Ouro Branco foi descontinuado pela instituição em 2019. Nota-se maior variação quando considerada a modalidade dos cursos, ao contrário do encontrado na tabela anterior.

Tabela 8: Cinco cursos com menos solicitações de bolsas

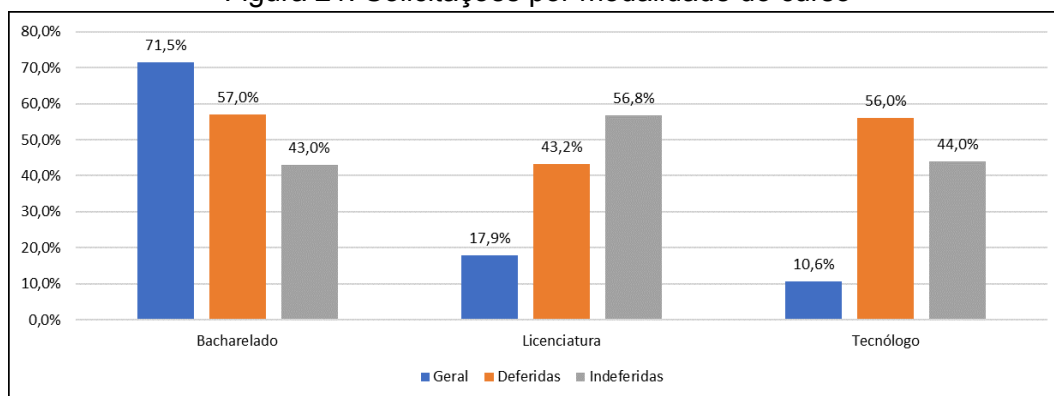
Campus	Curso	Modalidade	Deferidas	Indeferidas	Total
Bambuí	Educação Física	Licenciatura	3 (75,0%)	1 (25,0%)	4
Ouro Branco	Computação	Licenciatura	5 (38,5%)	8 (61,5%)	13
Ouro Preto	Gastronomia	Tecnólogo	17 (54,8%)	14 (45,2%)	31
Sabará	Administração	Bacharelado	21 (53,8%)	18 (46,2%)	39
Ponte Nova	Processos Gerenciais	Tecnólogo	33 (80,5%)	8 (19,5%)	41

Fonte: Elaborada pelo autor (2023)

Os cursos de Bacharelado possuem a maior quantidade de solicitações de bolsas, seguidos pelos cursos de Licenciatura e Tecnólogo (Figura 21). Quando comparados os resultados das solicitações, os cursos de Bacharelado e Tecnólogo possuem indicadores próximos de deferimento, ao contrário dos cursos de Licenciatura, cujo percentual de indeferimento é maior. Tal resultado pode indicar a necessidade de compreender melhor as razões pelas quais os candidatos vinculados a cursos de Licenciatura não estão obtendo êxito nas suas requisições de bolsas, fator que, eventualmente, pode estar impactando nas taxas de evasão desses cursos. A maioria dos cursos ofertados no IFMG são da modalidade Bacharelado, seguido pelas Licenciaturas e Tecnólogos.

Os gráficos das Figuras 21, 22, 23 e 24 compartilham a mesma escala pela necessidade de representar, simultaneamente, os percentuais gerais de solicitações entre as categorias do atributo analisado (barras azuis) e os percentuais de solicitações deferidas e indeferidas dentro de uma mesma categoria do atributo (barras laranjas e cinzas). Embora essas duas variáveis estejam expressas em termos percentuais, elas se referem a proporções distintas e com significados analíticos diferentes: a primeira diz respeito à participação relativa de determinada categoria no total de solicitações, enquanto a segunda expressa a taxa de deferimento (ou indeferimento) dentro de uma mesma categoria.

Figura 21: Solicitações por modalidade do curso

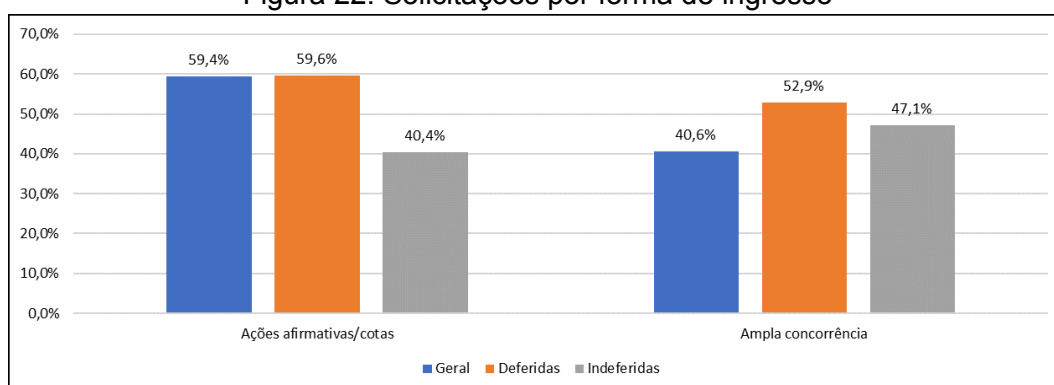


Fonte: Elaborada pelo autor (2023)

Os cursos do IFMG são ofertados em diferentes turnos: diurno (matutino ou vespertino), noturno ou integral. 54% das solicitações de bolsas foram feitas por candidatos matriculados em cursos noturnos, 38% em cursos de período integral e apenas 8% em cursos conduzidos em um período do dia. Embora exista diferença no quantitativo de solicitações entre os turnos de oferta dos cursos, os percentuais de deferimento e indeferimento são similares, em torno de 57% dos pedidos foram atendidos e 43% não atendidos. Esta pouca variação entre os deferidos e indeferidos de categorias diferentes também foi observada com as solicitações agrupadas conforme a área do curso na CAPES. A maior parte das solicitações envolveu candidatos de cursos das Engenharias (34%), seguido das Ciências Sociais Aplicadas (23%), Ciências Agrárias (18%), Ciências Exatas e da Terra (14%), Linguística, Letras e Artes (5%) e o restante em outras áreas.

Quando analisadas as solicitações pela forma de ingresso do candidato na instituição, percebe-se na Figura 22 que a maioria (59%) advém de estudantes que utilizaram o sistema de ações afirmativas e cotas no processo seletivo, enquanto o restante (41%) ingressou por ampla concorrência. Embora os resultados das solicitações sejam relativamente próximos entre esses dois grupos, ambos com percentual de deferimento superior a 50%, houve mais deferimentos de solicitações feitas por ingressantes sob ações afirmativas e cotas, indicando maior abrangência de atendimento a um público com demanda existente desde antes da entrada no curso superior.

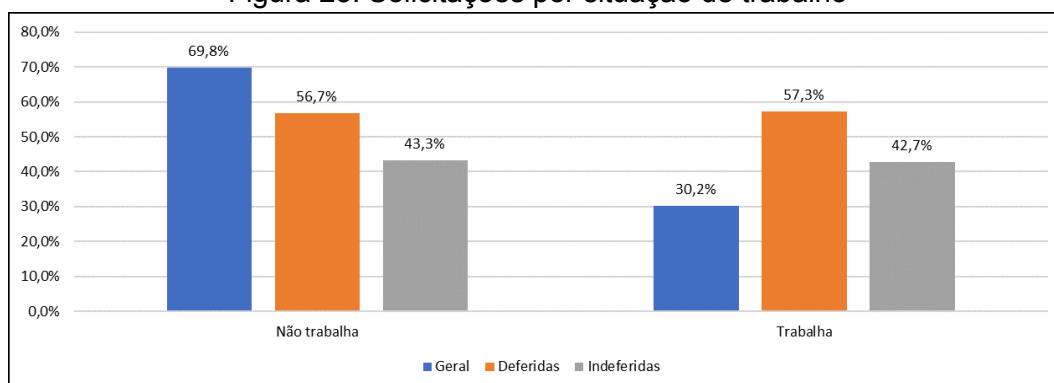
Figura 22: Solicitações por forma de ingresso



Fonte: Elaborada pelo autor (2023)

O fato de o candidato trabalhar no momento da solicitação de auxílio socioeconômico não possui influência direta no deferimento do benefício, conforme observa-se na Figura 23. Embora 70% das solicitações tenham sido feitas por candidatos que não trabalhavam, o que, por sua vez, pode indicar sob certo ponto de vista uma necessidade maior de obtenção do auxílio, contra 30% de solicitações de candidatos que trabalhavam, a variação entre os percentuais de deferimento e indeferimento nesses dois grupos é muito baixa, sugerindo que este atributo não tenha significativo impacto como fator para determinação da concessão da bolsa.

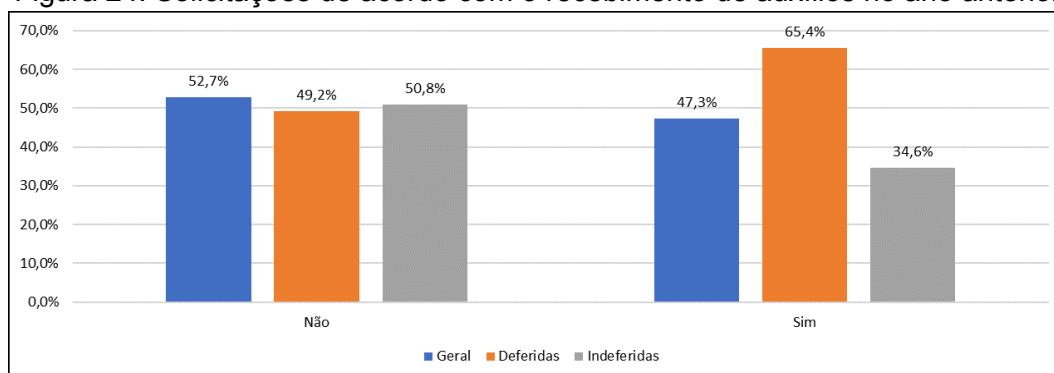
Figura 23: Solicitações por situação de trabalho



Fonte: Elaborada pelo autor (2023)

O recebimento de auxílio socioeconômico no ano anterior aparenta ser um importante atributo no que tange à admissão do mesmo como bolsista, como apresentado na Figura 24. Apesar de a maioria das solicitações (52,7%) ter sido realizada por candidatos que não tinham recebido auxílio, o percentual de deferimento entre os candidatos do grupo que já havia recebido o auxílio foi mais de 15% superior. Além da manutenção de seu perfil de vulnerabilidade de um ano para o outro, o conhecimento sobre as nuances do processo seletivo da assistência estudantil pode ser outra razão capaz de explicar a diferença.

Figura 24: Solicitações de acordo com o recebimento de auxílios no ano anterior



Fonte: Elaborada pelo autor (2023)

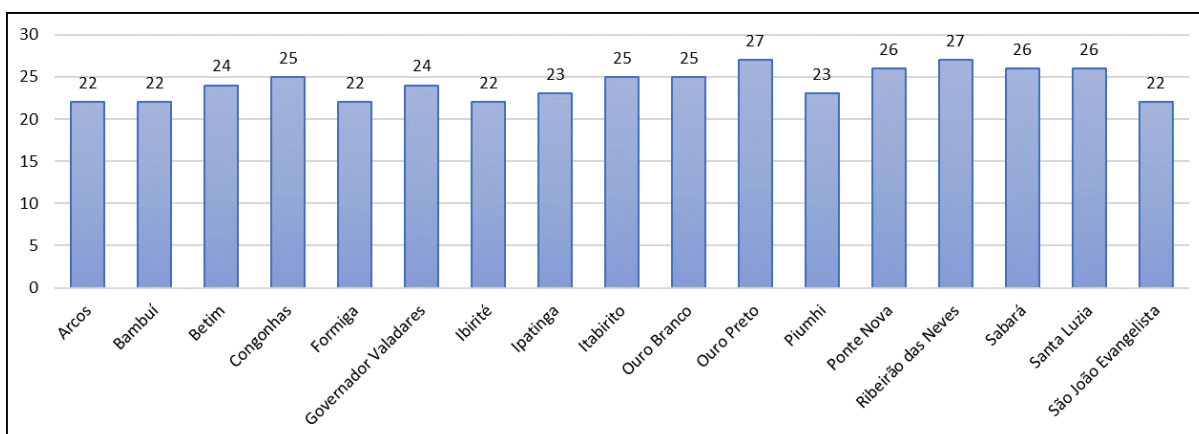
Alguns atributos do conjunto de dados apresentaram pouca variação em seus valores. Sobre o estado civil dos candidatos, 91% são solteiros, 7% casados e 2% com outras condições. 98% dos candidatos informaram não possuir qualquer tipo de deficiência, embora 14% tenham mencionado que possuem membros do grupo familiar com algum tipo de problema de saúde. O mesmo percentual (98%) de candidatos informou ter cursado o Ensino Fundamental e o Ensino Médio inteiramente em uma instituição pública. Pouco mais de 93% dos candidatos nasceram em Minas Gerais. 97% dos candidatos não receberam bolsa de pesquisa ou extensão antes de fazerem suas solicitações, enquanto 99% mencionaram que não receberam bolsa de monitoria ou tutoria.

Outros atributos coletados permitem caracterizar aspectos mais específicos dos candidatos e de seu grupo familiar. 56,9% dos candidatos indicaram “Feminino” como sua identidade de gênero, 42,8% indicaram “Masculino” e 0,3% indicaram a opção “Outra”. Quanto à cor, 49,8% se declararam pardos, 34% brancos, 14,8% pretos e 1,4% amarelos/indígenas. A maioria dos candidatos (60%) afirmou morar com familiares, 23% em repúblicas estudantis ou moradia institucional e 17% moram sozinhos. Sobre a situação do imóvel onde residem, 54% dos candidatos moram em imóvel próprio, 25% em locais alugados e 21% em imóveis com outras condições. 75% dos candidatos indicaram que não possuem imóveis além da moradia (nem o seu grupo familiar), mesmo percentual daqueles que afirmaram não serem os responsáveis pelo seu grupo familiar. O percentual de candidatos (ou algum membro do seu grupo familiar) que têm a propriedade de veículos é de 51%. Em relação ao meio utilizado para deslocamento até o Campus, a maioria (62%) utiliza algum tipo de transporte público, 14% utiliza transporte próprio e 24% não utiliza transporte ou se desloca de formas variadas, como carona, por exemplo.

Foi encontrada variação entre os *campi* na idade média dos candidatos no momento da inscrição no processo seletivo da assistência estudantil (Figura 25). Em Arcos, Bambuí, Formiga, Ibitiré e São João Evangelista a média de idade é a mais baixa (22 anos), enquanto a mais alta advém dos candidatos dos *campi* Ouro Preto e Ribeirão das Neves (27

anos). Dada esta diferença de cinco anos, esta informação pode ser útil para a gestão institucional por permitir que, eventualmente, sejam criados na política assistencial dispositivos específicos para candidatos em determinada faixa etária.

Figura 25: Idade média na inscrição por Campus



Fonte: Elaborada pelo autor (2023)

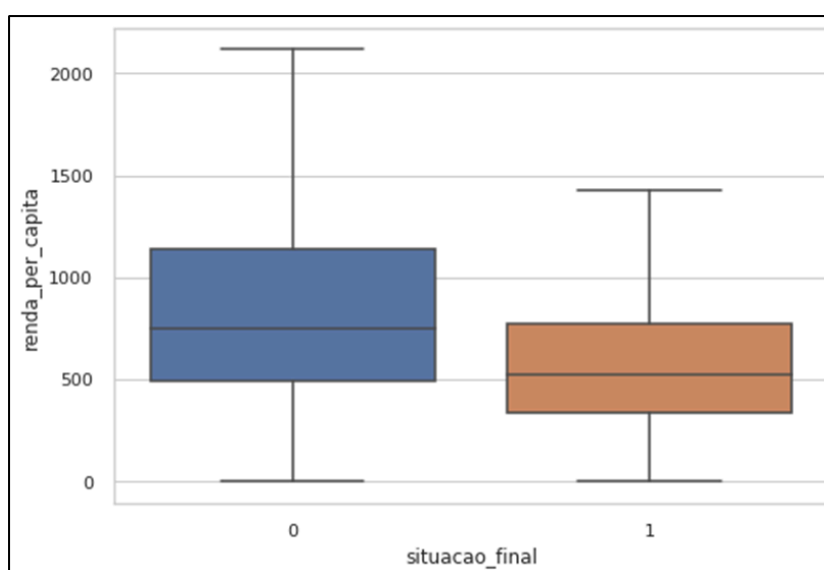
Para complementar o entendimento sobre os dados disponíveis, a técnica *Random Forest Feature Importance* (disponível na biblioteca do *scikit-learn*) foi executada visando identificar qual a importância de cada atributo para determinar o deferimento da solicitação de auxílio. O resultado com os atributos com importância mínima de 4% está disposto na Tabela 9. Observa-se que os atributos *renda_per_capita* e *idade_na_inscricao* respondem por quase a metade da importância geral, com 31% e 17%, respectivamente. Este achado mostra, de forma diferente da explorada no início desta Seção de caracterização dos candidatos a auxílio socioeconômico, que a renda possui grande influência sobre o resultado das solicitações, em consonância com o critério mínimo preconizado pelo PNAES. A idade do candidato no momento da solicitação também possui interessante capacidade preditora, mesmo não constando formalmente como um critério na política da instituição para a concessão do auxílio. Conforme mencionado anteriormente, o fato do candidato ter recebido o auxílio no ano anterior também contribui, mesmo que em menor proporção, para um resultado positivo. Analisando os outros atributos listados, destaque para a participação de dois atributos criados para este estudo, ambos com 5% de importância: *area_curso* e *turno_curso*.

Tabela 9: Importância dos atributos

Atributo	Importância	Atributo criado?
renda_per_capita	31%	Não
idade_na_inscricao	17%	Sim
cor	7%	Não
recebeu_auxilio_ano_anterior	6%	Não
num_pessoas_gf	5%	Não
area_curso	5%	Sim
turno_curso	5%	Sim
possui_veiculo	4%	Não
forma_ingresso	4%	Não

Fonte: Elaborada pelo autor (2023)

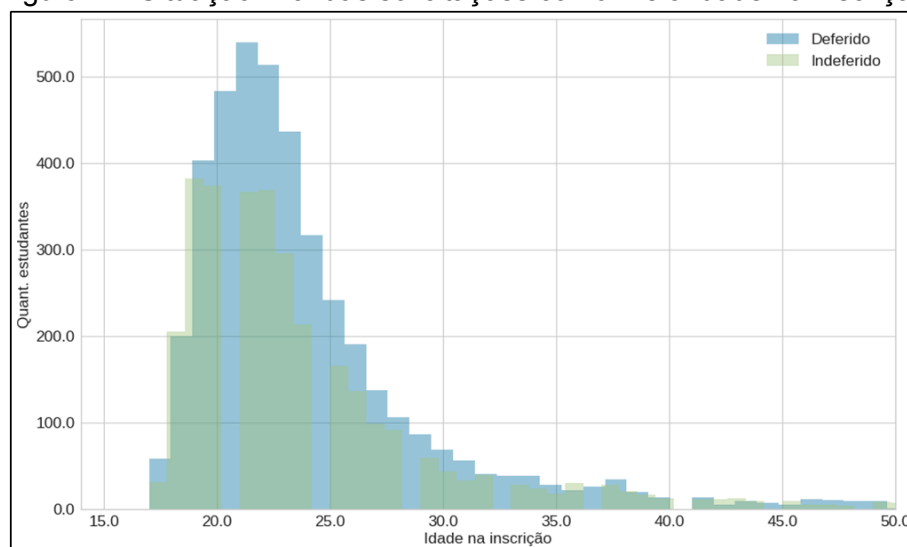
A Figura 26 provê outra forma de visualizar a importância da renda *per capita* do candidato para o deferimento de sua solicitação. É possível perceber que a renda média das solicitações indeferidas (primeira coluna, com o valor zero) é superior àquela das solicitações deferidas (segunda coluna, valor um), permitindo inferir que uma solicitação de auxílio advinda de um candidato com renda *per capita* baixa possui maior chance de deferimento.

Figura 26: Renda *per capita* de acordo com a situação final das solicitações

Fonte: Elaborada pelo autor (2023)

Quanto à idade na inscrição do processo seletivo, a Figura 27 mostra a distribuição da idade dos candidatos em função do resultado final. Observa-se maior diferença entre o volume de deferimentos e indeferimentos no intervalo de idade de 20 a 25 anos, inclusive.

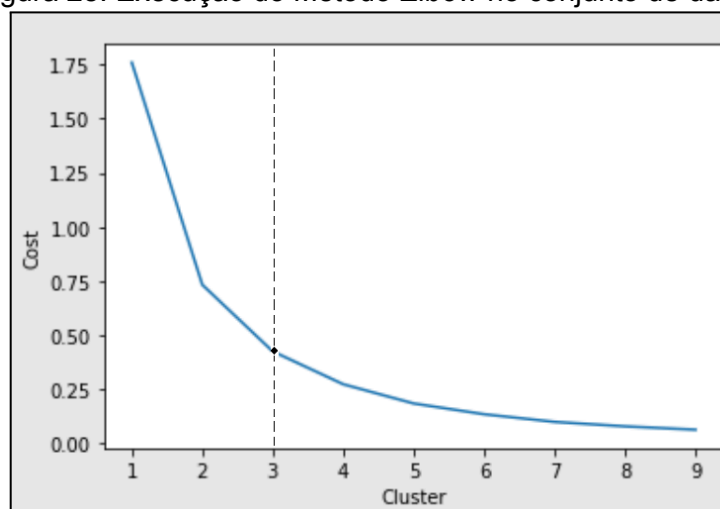
Figura 27: Situação final das solicitações conforme a idade na inscrição



Fonte: Elaborada pelo autor (2023)

A quarta atividade utilizou recursos de Aprendizado de Máquina para criar agrupamentos dos estudantes que participaram dos processos seletivos da Assistência Estudantil com base no grau de semelhança de seus atributos. Baseada no algoritmo *K-means*, o qual demanda a indicação de quantos agrupamentos devem ser criados logo na sua inicialização, a primeira tarefa consistiu na identificação desse número por meio da aplicação do Método *Elbow*. Na Figura 28 é possível notar que a quantidade ideal de agrupamentos inferida pelo método é igual a três, valor utilizado na sequência da atividade.

Figura 28: Execução do Método *Elbow* no conjunto de dados



Fonte: Elaborada pelo autor (2023)

Como o conjunto de dados possuía alta dimensionalidade, isto é, grande quantidade de atributos, fato que poderia dificultar o processo de visualização dos dados dos agrupamentos, optou-se pela aplicação da técnica *Principal Component Analysis* (PCA) para condensar estatisticamente os atributos em dois componentes denominados PC-1 e PC-2. O Quadro 1 apresenta os cinco atributos que mais contribuíram para a formação de cada um desses componentes.

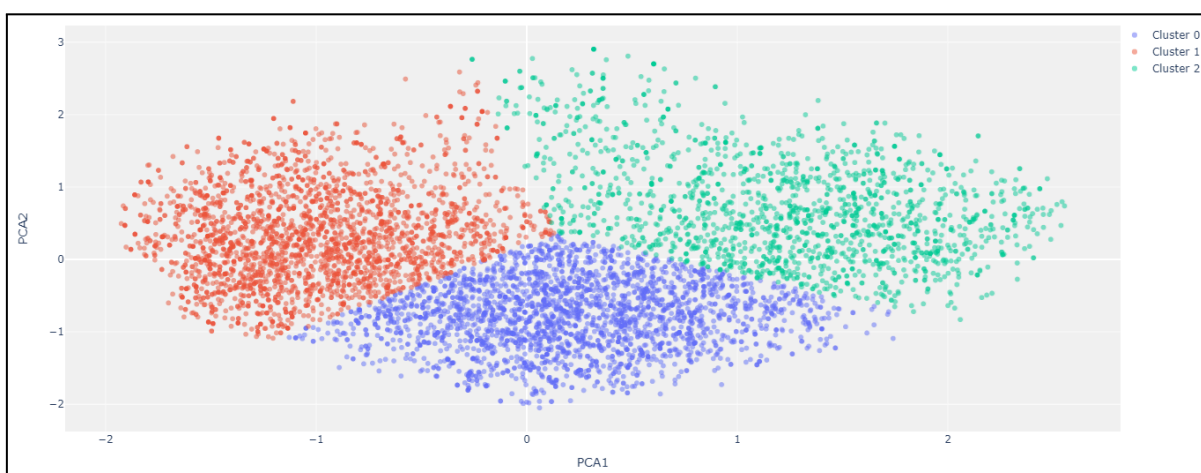
Quadro 1: Cinco atributos mais importantes de cada componente

PC-1	PC-2
area_curso	situacao_moradia
renda_per_capita	idade_na_inscricao
turno_curso	modalidade_curso
recebeu_auxilio_ano_anterior	cor
meio_transporte_para_campus	identidade_genero

Fonte: Elaborada pelo autor (2023)

Após a configuração do PCA, o algoritmo *K-means* foi executado com os dados normalizados e dividiu as instâncias em três agrupamentos numerados com 0, 1 e 2. Com o esquema de cores utilizado na Figura 29 pode-se visualizar o espaço ocupado pelos agrupamentos e a separação entre eles, bem como o volume de elementos que os compõem.

Figura 29: Agrupamentos criados pelo *K-means*

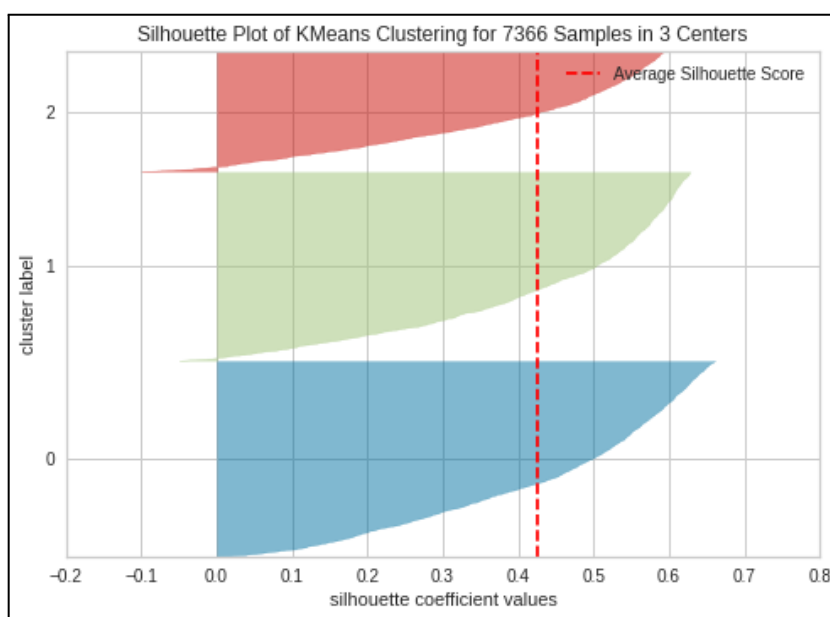


Fonte: Elaborada pelo autor (2023)

Com o intuito de avaliar se os elementos pertencentes a cada agrupamento são similares e, ao mesmo tempo, diferentes daqueles dos outros agrupamentos, o que

reforçaria a qualidade do modelo de agrupamento criado, a métrica *Silhouette Score* foi coletada e serviu de base para a elaboração da Figura 30. Embora a média geral da métrica tenha sido de 0,425 (em uma escala que vai de -1 a 1), nos três agrupamentos muitas instâncias tiveram resultados acima de 0,5. É plausível afirmar, portanto, que o modelo criado contém fronteiras bem definidas entre seus agrupamentos, além de elementos corretamente alocados em cada agrupamento.

Figura 30: Resultados da métrica *Silhouette Score*



Fonte: Elaborada pelo autor (2023)

A proposta da última atividade desta fase era promover a caracterização de cada um dos agrupamentos criados quanto aos aspectos demográficos e socioeconômicos de seus elementos, de maneira a permitir a diferenciação dos três perfis identificados pela estratégia de Aprendizado de Máquina adotada. O Quadro 2 foi elaborado com o objetivo de expor as características que se mostraram diferentes entre os agrupamentos, habilitando também a comparação entre eles. É válido ressaltar que foram exploradas no resultado final apenas características que possuíam, originalmente, maior variância entre os valores do conjunto de dados.

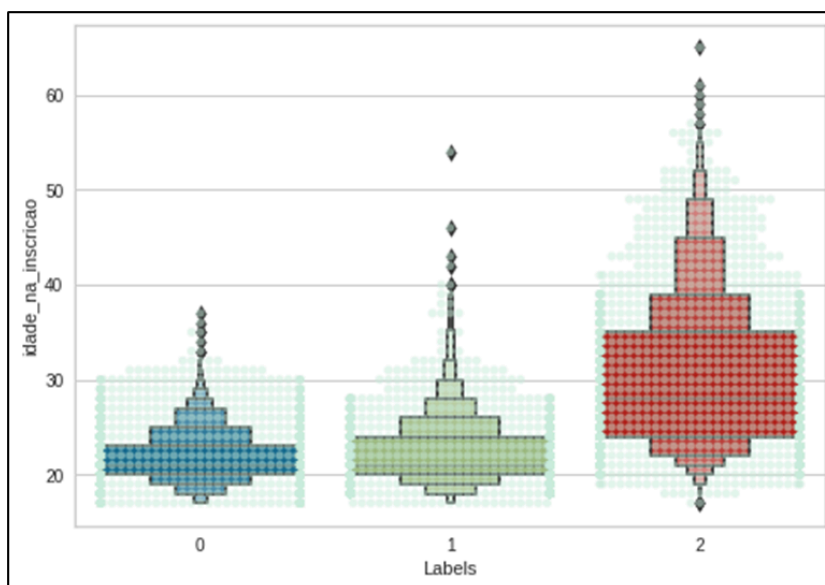
Quadro 2: Caracterização dos agrupamentos

Característica	Agrupamento 0	Agrupamento 1	Agrupamento 2
Quantidade de instâncias	2.856	2.760	1.750
Modalidade do curso	Equilíbrio entre Bacharelado, Licenciatura e Tecnólogo	Predomínio de Bacharelado	Equilíbrio entre Bacharelado, Licenciatura e Tecnólogo
Turno do curso	Predomínio de Noturno	Predomínio de Integral	Predomínio de Noturno
Área do curso	Predomínio de Ciências Sociais Aplicadas e Engenharias	Predomínio de Ciências Agrárias e Engenharias	Equilíbrio entre todas as áreas
Idade média na inscrição	21	22	28
Renda <i>per capita</i> média	R\$450,07	R\$1.100,00	R\$649,75
Identidade de gênero	Predomínio de Feminino	Predomínio de Masculino	Equilíbrio entre Feminino e Masculino
Situação do trabalho	Predomínio de candidatos que não trabalham	Predomínio de candidatos que não trabalham	Predomínio de candidatos que trabalham
Situação da moradia	Predomínio de candidatos que moram com familiares	Equilíbrio entre candidatos que moram com familiares e em república/moradia estudantil	Predomínio de candidatos que moram sozinhos
Situação do imóvel	Predomínio de candidatos que vivem em imóvel próprio	Predomínio de candidatos que vivem em imóvel próprio	Predomínio de candidatos que vivem em imóvel alugado
Possui veículo	Predomínio de candidatos que possuem veículo	Predomínio de candidatos que possuem veículo	Predomínio de candidatos que não possuem veículo
Responsável pelo Grupo Familiar (GF)	Predomínio de candidatos que não são responsáveis	Predomínio de candidatos que não são responsáveis	Predomínio de candidatos que são responsáveis
Recebeu auxílio no ano anterior	Predomínio de candidatos que não receberam	Equilíbrio entre beneficiários e não beneficiários	Equilíbrio entre beneficiários e não beneficiários
Número médio de pessoas no GF	2	3	3
Principal bolsa atribuída	BP4	BP4	BP4

Fonte: Elaborado pelo autor (2023)

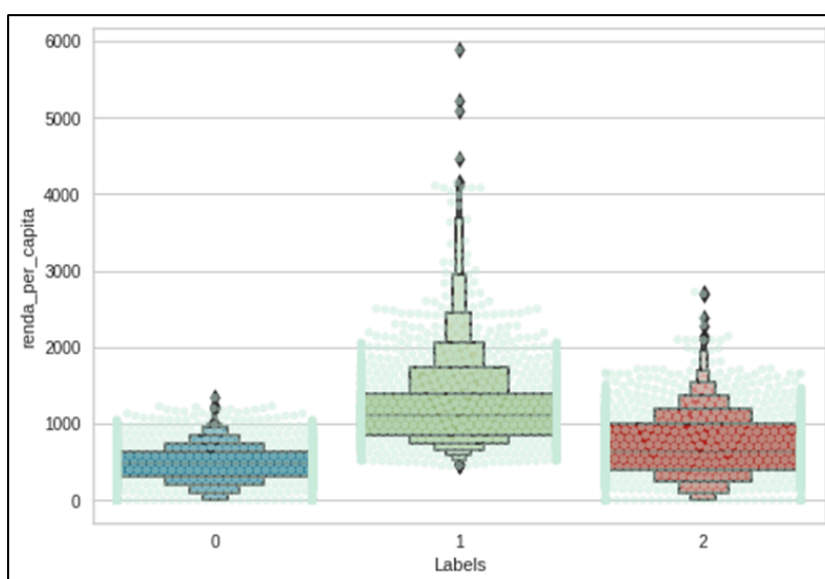
É interessante perceber as diferenças nos agrupamentos considerando os dois atributos identificados anteriormente como mais importantes para o deferimento das solicitações de auxílio: idade na inscrição e renda *per capita*. Observando a Figura 31, nota-se que a média de idade dos candidatos no momento da inscrição era mais baixa no primeiro e segundo agrupamentos (21 e 22 anos, respectivamente), enquanto no terceiro ficaram concentrados aqueles candidatos com maior idade (média de 28 anos).

Figura 31: Idade na inscrição nos três agrupamentos



Fonte: Elaborada pelo autor (2023)

No caso da renda *per capita*, cuja distribuição está representada na Figura 32, as médias finais tiveram maior variação entre os agrupamentos. No primeiro agrupamento encontram-se candidatos com as menores rendas, sendo que a instância com o maior valor sequer ultrapassa os R\$1.500,00. Já o terceiro agrupamento fica em uma posição intermediária, possuindo desde rendas baixas até instâncias com renda próxima de R\$3.000,00. No segundo agrupamento estão as instâncias com maiores rendas, sendo seu valor mínimo próximo da média geral do primeiro agrupamento.

Figura 32: Renda *per capita* nos três agrupamentos

Fonte: Elaborada pelo autor (2023)

O perfil dos candidatos do primeiro agrupamento indica que a maioria estuda à noite em cursos de todas as modalidades e vinculados com as áreas de Ciências Sociais Aplicadas e Engenharias; a idade média no momento da inscrição é baixa, assim como a renda *per capita*; feminino é a identidade de gênero predominante, sendo que a maioria não trabalha, reside com familiares em imóvel próprio, possui veículo e não recebeu auxílio no ano anterior; o grupo familiar é composto por, em média, duas pessoas.

Algumas características são bem diferentes no segundo agrupamento, composto por candidatos matriculados em cursos de Bacharelado das áreas de Ciências Agrárias e Engenharias, cujas atividades letivas ocorrem em período integral; embora a idade na inscrição também seja baixa como no primeiro agrupamento, a renda *per capita* média do segundo agrupamento é a mais alta; a maioria dos candidatos é do gênero masculino, não trabalha e não é responsável pelo grupo familiar, o qual é composto por 3 pessoas, em média; existe um equilíbrio entre candidatos que moram com familiares e em república ou moradia estudantil, assim como observa-se entre os candidatos que receberam e não receberam auxílio no ano anterior.

O terceiro agrupamento possui particularidades interessantes, como uma idade média na inscrição consideravelmente acima dos outros agrupamentos, nivelamento maior quanto às áreas dos cursos e gênero e o fato da maioria trabalhar; a renda *per capita* está em uma posição intermediária entre os agrupamentos; os candidatos estão matriculados em cursos de todas as modalidades, porém existe um predomínio de oferta no turno noturno; a maioria dos candidatos mora sozinha em imóvel alugado e não possui veículo, com média de três pessoas compondo o grupo familiar.

Ao final desta fase, é pertinente afirmar que a construção do perfil dos candidatos a auxílio socioeconômico do IFMG contribuiu diretamente para a execução dos outros objetivos propostos na tese, bem como poderá servir de suporte para o planejamento e tomada de decisões da instituição que envolvam o aprimoramento das políticas e práticas da assistência estudantil. A adequada distribuição dos recursos financeiros para os candidatos passa, primeiramente, pelo entendimento das características do público demandante, de maneira que o auxílio concedido realmente contribua para a manutenção dos estudantes em seus cursos com o devido desempenho.

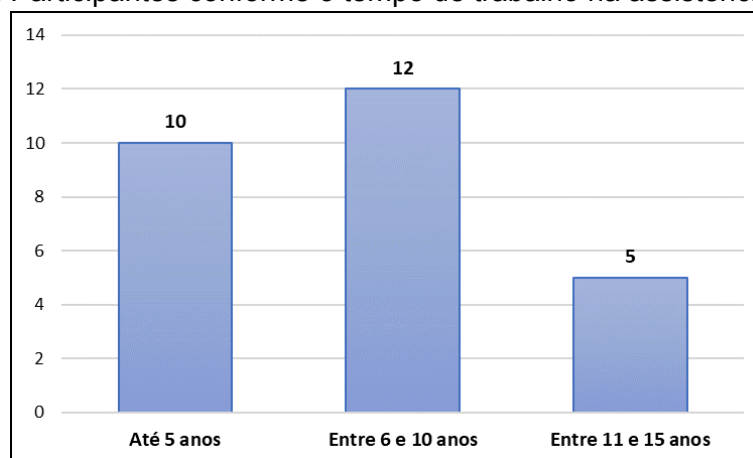
4.2 Caracterização do perfil de comportamento informacional dos servidores envolvidos com o processo de concessão de auxílio socioeconômico

O questionário para a coleta dos dados necessários à construção do perfil desejado nesta fase do estudo foi enviado via e-mail para os setores de assistência estudantil de todos os *campi* do IFMG. Foi solicitado ao responsável que encaminhasse o link de acesso ao formulário para os servidores que participaram de processos seletivos para distribuição

de auxílios socioeconômicos entre 2019 e 2022. Decorrido o prazo de quinze dias, 27 participações foram registradas, sendo que pelo menos um servidor de cada Campus respondeu o questionário. As respostas foram tabuladas por questão, permitindo identificar a frequência das opções escolhidas pelos participantes em cada uma delas. É importante lembrar que o questionário utilizado abrange as quatro categorias de contexto que perfazem o Comportamento Informacional: necessidade, busca, uso e compartilhamento da informação ([WILSON, 1999](#)).

A primeira parte do questionário visa caracterizar brevemente o perfil dos respondentes. Quanto à identidade de gênero, 22 participantes (81%) escolheram a opção “Feminino”, enquanto 5 (19%) escolheram “Masculino”. A média de idade dos participantes é de 36 anos, sendo que o mais jovem possui 25 anos e o mais velho 46 anos. A questão sobre o tempo de vínculo com a assistência estudantil mostrou que a maioria dos participantes já atua neste setor entre seis e dez anos, conforme observa-se na Figura 33. Servidores que atuam até cinco anos vêm na sequência, enquanto a menor parte dos respondentes trabalha no setor há mais de 11 anos.

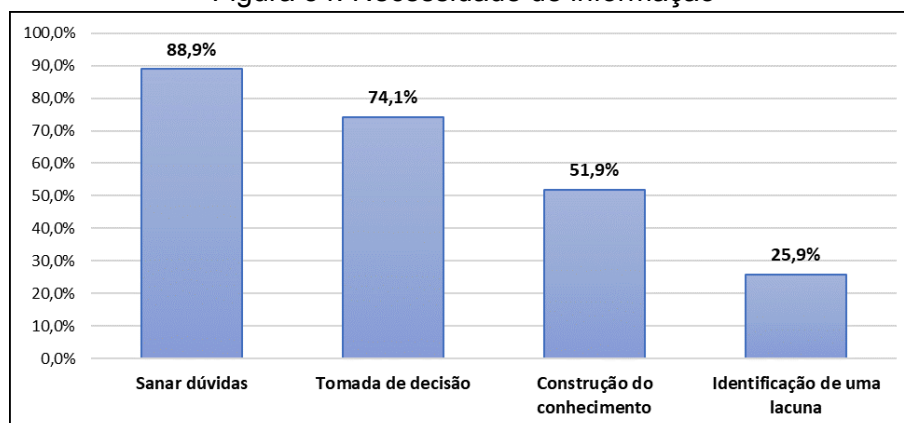
Figura 33: Participantes conforme o tempo de trabalho na assistência estudantil



Fonte: Elaborada pelo autor (2023)

A seção seguinte do questionário aborda aspectos relacionados com as quatro categorias do Comportamento Informacional. Inicialmente, foi analisada a necessidade de informação dos servidores participantes dos processos seletivos da assistência estudantil do IFMG. Utilizando a frequência de escolha das opções “Frequente” e “Muito frequente”, os resultados expostos na Figura 34 indicam que a principal necessidade informacional dos participantes se refere à busca por sanar dúvidas, sejam elas próprias ou de terceiros. A tomada de decisão também se mostra como um importante fator que leva os participantes a demandarem informações, permitindo relacionar esses dois motivos com o trabalho em si dos servidores deste setor, em especial quando considerada a natureza decisória envolvida.

Figura 34: Necessidade de informação



Fonte: Elaborada pelo autor (2023)

O comportamento de busca da informação foi analisado com base em quatro subcategorias que, por sua vez, abordam aspectos advindos do processo de pesquisa: (I) tipos de informação necessários para a execução do trabalho dos respondentes; (II) fontes e os canais de comunicação; (III) barreiras enfrentadas; (IV) sentimentos experimentados. A Tabela 10 apresenta os resultados da questão responsável por indicar os tipos de informações demandadas. Nota-se que o item “Perfil socioeconômico do estudante” foi citado como “Frequente” ou “Muito frequente” por todos os participantes. Além disso, os quatro itens seguintes também possuem considerável importância no processo de busca, pois fazem parte das rotinas cotidianas dos servidores do setor de assistência estudantil, em especial para os processos seletivos de bolsistas. Por exemplo, informações orçamentárias são fundamentais para dimensionar o quantitativo de estudantes a serem atendidos pelas políticas institucionais, bem como informações sobre o aproveitamento acadêmico e evasão dos estudantes fundamentam eventuais tomadas de decisão sobre a atribuição de auxílios. Por outro lado, percebe-se que informações vinculadas com os três últimos itens não fazem parte do contexto experimentado pelos respondentes.

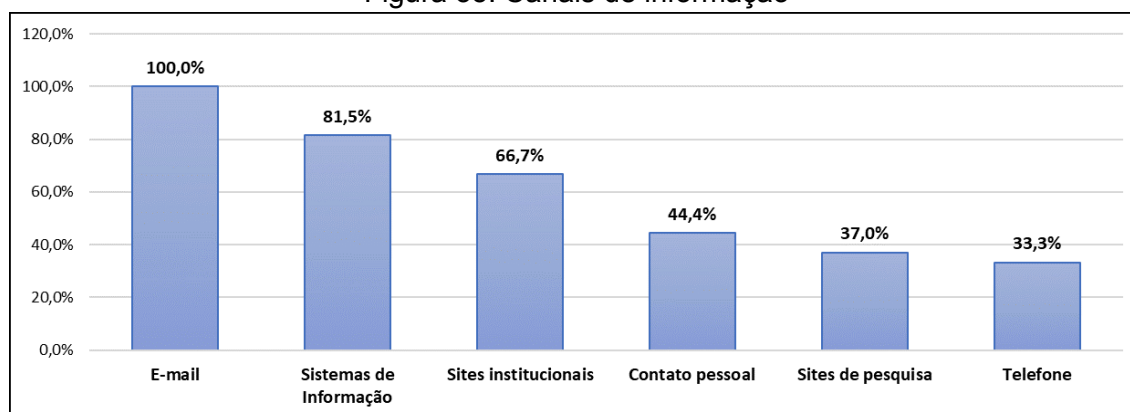
Tabela 10: Tipos de informação

Tipo de informação	Nunca	Raramente	Ocasional-mente	Frequente	Muito frequente	Frequente e muito frequente
Perfil socioeconômico do estudante	0,0%	0,0%	0,0%	33,3%	66,7%	100%
Acompanhamento pedagógico e/ou social dos estudantes	0,0%	7,4%	7,4%	63,0%	22,2%	85,2%
Aproveitamento acadêmico dos estudantes	11,1%	11,1%	18,5%	48,1%	11,1%	59,3%
Orçamento e Finanças	7,4%	7,4%	25,9%	37,0%	22,2%	59,3%
Evasão estudantil	14,8%	7,4%	18,5%	51,9%	7,4%	59,3%
Estudantes com necessidades educacionais específicas	7,4%	18,5%	37,0%	33,3%	3,7%	37,0%
Refeições servidas pelo Restaurante Universitário	81,5%	3,7%	0,0%	14,8%	0,0%	14,8%
Processos Licitatórios	51,9%	22,2%	22,2%	3,7%	0,0%	3,7%
Contratação de pessoal	63,0%	11,1%	25,9%	0,0%	0,0%	0,0%
Capacitação dos servidores	81,5%	11,1%	7,4%	0,0%	0,0%	0,0%

Fonte: Elaborada pelo autor (2023)

A questão relacionada com os canais de informação visa identificar por quais meios as informações chegaram até os participantes. Utilizando o percentual de respostas que marcaram “Frequente” e “Muito frequente” para cada item, a Figura 35 relaciona os seis veículos com melhor resultado. Merece destaque a alta importância atribuída pelos participantes aos três primeiros itens, indicando forte emprego de meios eletrônicos como canal para busca de informações necessárias ao desenvolvimento das atividades do setor de assistência estudantil.

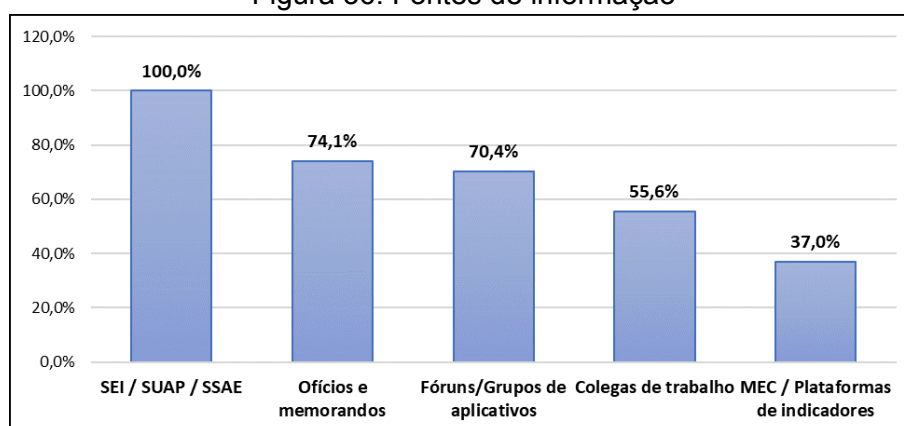
Figura 35: Canais de informação



Fonte: Elaborada pelo autor (2023)

Análise similar pode ser feita com os resultados da questão envolvendo as fontes de informação (expostos na Figura 36). Todos os participantes indicaram alta frequência de emprego como fontes os sistemas de informação utilizados pela instituição, provavelmente pelo fato de os mesmos fazerem parte dos processos informacionais internos. Os ofícios e memorandos, assim como mensagens trocadas em grupos de aplicativos móveis, também servem como fontes auxiliares aos participantes. A menor frequência observada no item “Colegas de trabalho” talvez seja explicada pelo fato de os setores de assistência estudantil nos *campi* da instituição não possuírem muitos servidores alocados, fazendo com que todo o processo de busca de informações seja conduzido, muitas vezes, por apenas uma pessoa.

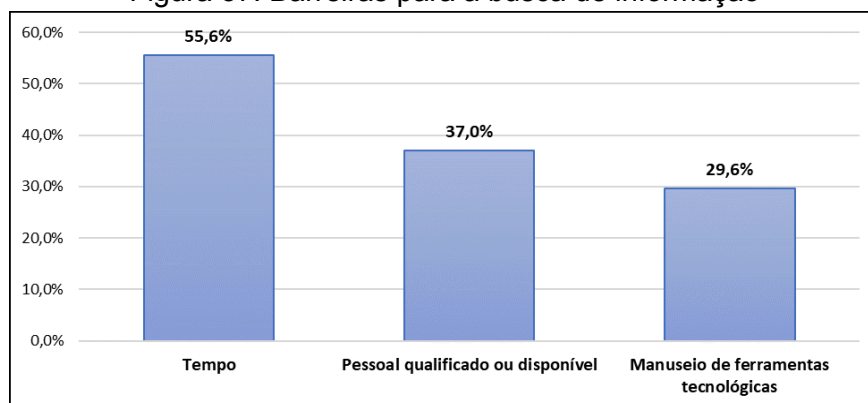
Figura 36: Fontes de informação



Fonte: Elaborada pelo autor (2023)

Os participantes apontaram algumas barreiras encontradas ao longo da busca de informações, as quais são caracterizadas por Wilson (1999) como variáveis intervenientes e sofrem influências de fatores pessoais, sociais e ambientais. Observa-se na Figura 37 que a barreira mais frequentemente indicada foi o tempo, seguida pela dificuldade em obter auxílio, no momento da busca, de pessoal qualificado e com disponibilidade. O manuseio de ferramentas tecnológicas também se mostrou como um fator influenciador no processo de busca com importância regular, situação que demanda uma análise mais detalhada por parte da instituição, haja vista a predominância do meio digital como canal e fonte de informação no contexto dos servidores da assistência estudantil.

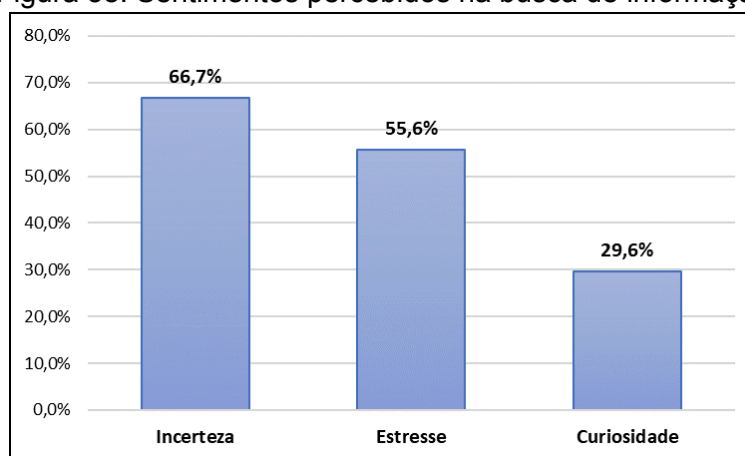
Figura 37: Barreiras para a busca de informação



Fonte: Elaborada pelo autor (2023)

O processo de busca de informações normalmente é construído não somente com as ações dos envolvidos, mas também com a investigação dos sentimentos percebidos por eles ao longo desta experimentação. Verifica-se na Figura 38 que 67% dos participantes indicaram a incerteza como o sentimento mais frequente durante as suas buscas por informações, seguido pelo estresse, caracterizando, de certa forma, um processo de busca desgastante.

Figura 38: Sentimentos percebidos na busca de informação



Fonte: Elaborada pelo autor (2023)

A terceira categoria do Comportamento Informacional levantada nesta pesquisa refere-se ao uso da informação por parte dos envolvidos, isto é, com qual propósito as informações buscadas são utilizadas, cujos resultados estão na Tabela 11. A maior parte dos participantes da pesquisa indicou que o principal motivo para uso das informações no contexto da assistência estudantil é servir como base para a elaboração de ações e metas para este setor, além de contribuírem para a elaboração de documentos formais de prestação de contas. A tomada de decisão também foi valorizada pelos respondentes como uma das razões para a continuidade dos processos de busca da informação e emprego dos resultados obtidos, possivelmente em face das questões relacionadas com a decisão de

atribuir os auxílios socioeconômicos. Por outro lado, fica bem claro ao se analisar os últimos itens o que não está diretamente vinculado com o escopo de atuação dos servidores da assistência estudantil.

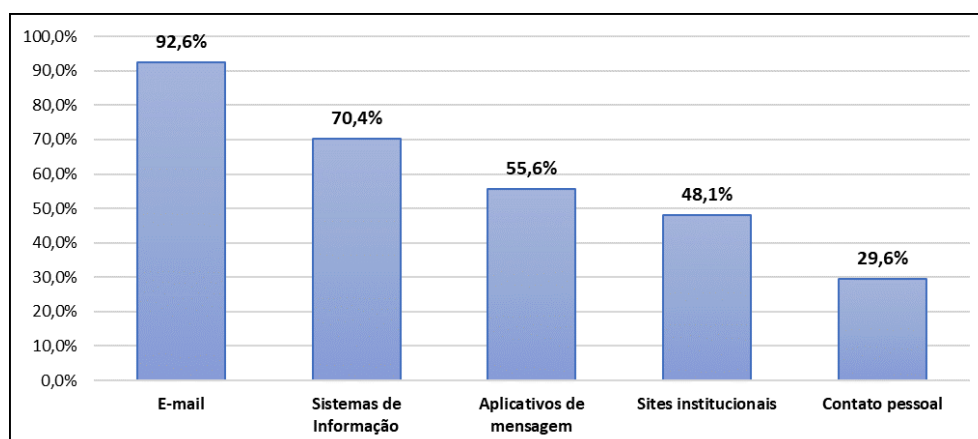
Tabela 11: Usos da informação

Uso da informação	Nunca	Raramente	Ocasional-mente	Frequente	Muito frequente	% frequente e muito frequente
Elaboração de ações/metad para assistência estudantil	3,7%	0,0%	3,7%	51,9%	40,7%	92,6%
Elaboração de relatórios para prestação de contas	0,0%	3,7%	11,1%	55,6%	29,6%	85,2%
Tomada de decisão	7,4%	11,1%	11,1%	37,0%	33,3%	70,4%
Elaboração de políticas afirmativas e de acessibilidade	3,7%	3,7%	51,9%	40,7%	0,0%	40,7%
Proposição de debates sobre temas específicos da área	37,0%	25,9%	18,5%	11,1%	7,4%	18,5%
Proposição de cursos e oficinas para estudantes e servidores	66,7%	3,7%	22,2%	3,7%	3,7%	7,4%
Composição da equipe de trabalho	51,9%	22,2%	22,2%	3,7%	0,0%	3,7%
Aquisição de bens e materiais	29,6%	18,5%	51,9%	0,0%	0,0%	0,0%
Construção de parcerias com outras instituições	63,0%	14,8%	22,2%	0,0%	0,0%	0,0%

Fonte: Elaborada pelo autor (2023)

O compartilhamento de informações foi analisado como sendo uma categoria do Comportamento Informacional e os resultados da pesquisa mostram, novamente, a preferência dos participantes pelos meios digitais em contraponto a outras formas de contato pessoal. É possível notar na Figura 39 que o e-mail e os sistemas informatizados são frequentemente utilizados como meio para compartilhar informações no contexto da instituição. Os aplicativos de mensagens também merecem destaque, sendo possivelmente um meio mais informal para a efetivação do compartilhamento.

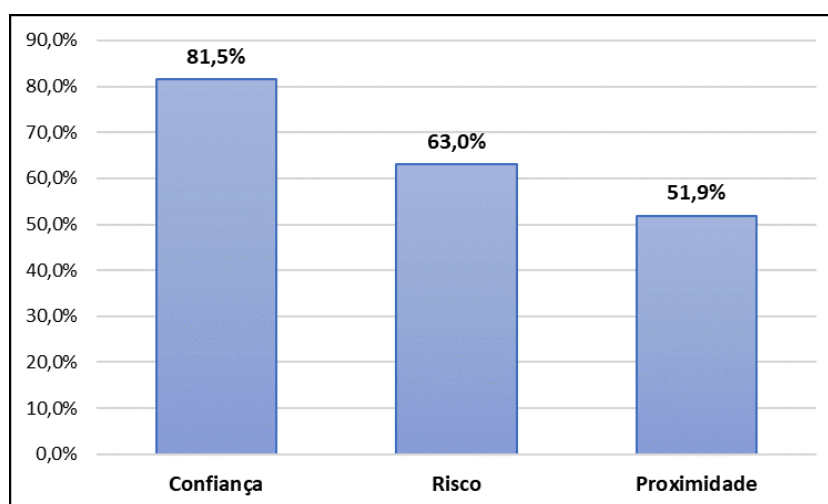
Figura 39: Meios de compartilhamento de informação



Fonte: Elaborada pelo autor (2023)

Ainda relacionado com o compartilhamento de informações, os participantes da pesquisa indicaram alguns elementos que consideram no momento desta ação. Observa-se na Figura 40 que a confiança no receptor das informações e no meio escolhido é o fator mais ponderado, seguido dos riscos assumidos ao realizar o compartilhamento de informações e da proximidade organizacional com o receptor.

Figura 40: Elementos considerados ao compartilhar informações



Fonte: Elaborada pelo autor (2023)

A última questão do instrumento de pesquisa objetivou coletar dados que permitissem compreender como os participantes caracterizam uma “boa informação” ou “informação valiosa”. De acordo com a Tabela 12, as três características mais importantes de uma informação para os servidores do setor de assistência estudantil do IFMG compreendem o fato dela ser completa (dispondo de todos os fatos significativos de acordo

com a busca empreendida), simples (permitindo uma interpretação clara e transparente do conteúdo) e precisa (sem erros que possam comprometer o seu entendimento). No outro extremo, como características com menor importância sob a percepção dos participantes, estão o fato de a informação ser pontual (ser entregue no prazo adequado), verificável (indicando a possibilidade de conferir sua correte) e flexível (para utilização com variados propósitos).

Tabela 12: Importância das características da informação

Característica	% Mais importante
Completa (contém todos os fatos relevantes)	70,4%
Simple (de fácil entendimento, sem muitos detalhes)	66,7%
Precisa (não contém erro)	66,7%
Confiável (os métodos de coleta e a origem são conhecidos)	63,0%
Econômica (não demandou custo excessivo na sua disponibilização)	55,6%
Acessível (de fácil obtenção por parte dos usuários autorizados)	55,6%
Segura (o acesso só é permitido aos usuários autorizados)	55,6%
Relevante (essencial para o processo de tomada de decisões)	51,9%
Flexível (poderá ser usada para uma variedade de propósitos)	44,4%
Verificável (possível de se conferir e assegurar que está correta)	44,4%
Pontual (disponibilizada dentro do prazo estipulado)	37,0%

Fonte: Elaborada pelo autor (2023)

Com o propósito de sumarizar os resultados obtidos na pesquisa, o Quadro 3 relaciona-os com o suporte teórico provido pelo Modelo de Comportamento Informacional de [Wilson \(1999\)](#). Ressalta-se que os itens de resultados apresentados são apenas aqueles que obtiveram, no mínimo, 50% de frequência quando escolhidas as opções “Frequente” e “Muito frequente”.

Quadro 3: Sumarização dos resultados sobre Comportamento Informacional

Categoria	Caracterização/subcategoria	Item(ns) de resposta
Necessidade de informação	Originadas a partir de necessidades internas e externas dos servidores da assistência estudantil	Sanar dúvidas
		Tomada de decisão
		Construção do conhecimento
Busca de informação	Tipos de informação	Perfil socioeconômico do estudante
		Acompanhamento pedagógico e/ou social dos estudantes
		Aproveitamento acadêmico dos estudantes
		Orçamento e Finanças
		Evasão estudantil
	Canais de informação	E-mail
		Sistemas de Informação
		Sites institucionais
	Fontes de informação	SEI/SUAP/SSAE
		Ofícios e memorandos
		Fórums/Grupos de aplicativos
		Colegas de trabalho
	Barreiras encontradas (variáveis intervenientes)	Tempo
	Sentimentos percebidos	Incerteza
		Estresse
Uso da informação	A utilização das informações obtidas ocorre tanto pelos servidores quanto para repasse a terceiros	Elaboração de ações/metadados para assistência estudantil
		Elaboração de relatórios para prestação de contas
		Tomada de decisão
Compartilhamento da informação	Meios de compartilhamento	E-mail
		Sistemas de Informação
		Aplicativos de mensagem
	Elementos considerados	Confiança
		Risco
		Proximidade

Fonte: Elaborado pelo autor (2023)

Diante do exposto, é possível concluir que, de maneira geral, o comportamento informacional dos servidores que participaram de processos seletivos da assistência estudantil do IFMG está diretamente associado ao escopo das suas tarefas no ambiente organizacional, cujas demandas vão desde a resolução de dúvidas próprias e de terceiros até a tomada de decisões diversas, em especial aquelas que envolvem a atribuição de auxílios socioeconômicos para estudantes com vulnerabilidade social. Quanto ao fato de que na busca e no compartilhamento de informações existe a preferência pelo emprego de soluções e ferramentas eletrônicas, em contraponto ao contato pessoal, é válido ressaltar que tal característica é benéfica sob o ponto de vista da formalização do fluxo informacional, mas, por outro lado, poderia ser interessante para a instituição fomentar o contato entre os servidores atuantes no setor com aqueles que detêm maior experiência. Outro ponto a ser evidenciado é que os principais tipos de informação requeridos pelos participantes da pesquisa em seu cotidiano estão alinhados com os construtos investigados nesta tese: auxílio socioeconômico, evasão escolar e desempenho acadêmico.

4.3 Construção e avaliação da ontologia de domínio

As atividades necessárias para a elaboração da ontologia de domínio seguiram o suporte metodológico provido pela SABiO, englobando as cinco fases de processos de desenvolvimento e alguns processos de apoio. Dentre os diversos artefatos produzidos, destacam-se a ontologia de referência fundamentada com a UFO e construída com a linguagem OntoUML, bem como a ontologia operacional que permitiu a organização e leitura dos dados no ambiente computacional construído sob o formato de um *Graph Database*.

É importante ressaltar que a ontologia de domínio resultante dessa etapa tem como escopo exclusivo o contexto do IFMG, baseado em um estudo de caso que guiou a identificação do problema, o levantamento dos requisitos e o desenvolvimento das soluções propostas. Os conceitos, classes e relações definidas na ontologia refletem diretamente os elementos observados no ambiente do IFMG, de modo a possibilitar aplicabilidade prática dos resultados obtidos com dados reais da instituição. Portanto, não é objetivo da tese expandir o modelo para abranger todo o domínio da educação superior ou da assistência estudantil, com vistas a se tornar uma *core ontology*. Embora tal delimitação possa ser interpretada como uma limitação à construção de uma ontologia mais ampla, ela não inviabiliza o reaproveitamento do artefato ontológico por instituições com realidades semelhantes ao caso aqui estudado, seja por meio da adoção direta do modelo ou com adaptações e extensões que contemplem as especificidades de cada instituição.

4.3.1 Identificação de propósito e definição de requisitos

O objetivo central desta fase era compreender o propósito da ontologia como elemento para organização e integração dos dados de desempenho acadêmico dos estudantes, evasão escolar e concessão de auxílio socioeconômico pela assistência estudantil do IFMG. Houve, ainda, a elicitación dos principais requisitos envolvidos neste contexto e o seu registro no formato de questões de competência, as quais serviram, posteriormente, para uma avaliação da qualidade da ontologia. Essas atividades foram conduzidas pelo pesquisador e tiveram como subsídios o amplo levantamento bibliográfico realizado desde o início do desenvolvimento desta tese, em conjunto com os resultados explorados nas seções 4.1 e 4.2 deste Capítulo.

A ontologia elaborada tem como propósito: “Recuperar informações sobre a distribuição de auxílio socioeconômico no formato de bolsa permanência pela assistência estudantil para analisar sua contribuição na redução da evasão e no aumento do desempenho acadêmico dos estudantes de cursos superiores presenciais do IFMG.”

As questões de competência são apresentadas a seguir e buscaram explorar, ao máximo, aspectos das relações entre os três construtos:

QC1: O desempenho acadêmico dos estudantes atendidos por auxílio socioeconômico é superior ao daqueles que não são atendidos?

QC2: Os estudantes que recebem auxílio socioeconômico evadem menos proporcionalmente do que aqueles que não recebem?

QC3: A proporção de estudantes formados é maior entre aqueles que recebem auxílio socioeconômico?

QC4: Existe relação entre a evasão, desempenho acadêmico e as características pessoais dos estudantes bolsistas e não bolsistas, tais como gênero e cor?

QC5: Existem variações no desempenho acadêmico e na taxa de evasão de acordo com o tipo de Bolsa Permanência destinada aos estudantes?

QC6: Como é a distribuição da evasão considerando as áreas de conhecimentos dos cursos, além da taxa de evasão de cada área?

QC7: Como é a distribuição da evasão considerando o grau certificado pelos cursos, além da taxa de evasão de cada grau?

QC8: Como é a distribuição da evasão considerando a forma de ingresso do estudante ao curso, além da taxa de evasão de cada forma?

O levantamento das questões de competência contribuiu para a organização de uma visão mais otimizada sobre como deveria ser a abrangência da ontologia dentro do contexto dos três construtos no IFMG, pois embora cada um deles tenha suas particularidades, era fundamental que o escopo da ontologia contemplasse tão somente aspectos necessários

para o alcance dos objetivos da pesquisa. Destaca-se, ainda, que as questões de competência definidas permitem tanto a exploração dos dados categorizados por bolsistas ou não bolsistas (questões 1 a 4) quanto análises mais abrangentes, usando para isso atributos caracterizadores do conjunto de dados, indicando a capacidade da ontologia em responder questões direcionadas ou não.

4.3.2 Captura e formalização da ontologia

Iniciando as atividades desta fase, o pesquisador analisou um documento fornecido pela Diretoria de Tecnologia da Informação (DTI) do IFMG contendo detalhes acerca da estrutura de armazenamento dos dados do sistema informatizado responsável pela gestão acadêmica da instituição. Por se tratar de documento interno do setor, a DTI solicitou que o mesmo não tivesse exposição direta nessa tese. Observou-se que, assim como no sistema informatizado da Assistência Estudantil, é utilizado um banco de dados relacional para gerenciar as estruturas de armazenamento de dados, sendo que o sistema acadêmico apresenta complexidade alta para entendimento das relações entre as tabelas, pois é um *Enterprise Resource Planning* (ERP) cujo escopo extrapola o domínio da gestão acadêmica.

Em conjunto com os documentos citados, o IFMG também disponibilizou ao pesquisador dois arquivos com os dados acadêmicos dos seus estudantes conforme o escopo abrangido por esta pesquisa, qual seja: cursos superiores presenciais de graduação, considerando as turmas de ingressantes entre o primeiro semestre letivo de 2018 e o segundo semestre letivo de 2022. O primeiro arquivo estava no formato de planilha eletrônica e continha os seguintes dados dos estudantes: *identificador do estudante, campus, curso, data de nascimento, gênero, naturalidade, cor, forma de ingresso, ano de encerramento do Ensino Médio, ano de ingresso no curso, ano de conclusão do curso, turno de estudos, carga horária total do curso, coeficiente de rendimento geral e situação atual da matrícula*. O segundo arquivo estava no formato CSV em função da grande quantidade de registros (superior a um milhão e quinhentos mil), dispendo os dados das disciplinas cursadas por cada estudante: *identificador do estudante, campus, curso, período letivo, nome da disciplina, nota final e situação final na disciplina*.

Merece registro o fato de o sistema acadêmico não armazenar tantos dados categóricos dos estudantes como acontece no da Assistência Estudantil, cuja quantidade de atributos permite caracterizar os estudantes sob diferentes enfoques, como realizado na Seção 4.1. Por outro lado, observa-se que o sistema acadêmico disponibiliza dois dados fundamentais para possibilitar o cumprimento do objetivo proposto nesta tese, o coeficiente de rendimento geral de cada estudante (para avaliação do desempenho acadêmico) e a situação atual da matrícula (indicando a ocorrência de evasão, permanência ou conclusão do curso). É válido ressaltar que a frequência escolar dos estudantes não foi disponibilizada

pela instituição, em virtude de se tratar de registros com baixo grau de confiança de preenchimento por parte dos docentes.

Com o intuito de identificar características que, eventualmente, poderiam ser contempladas na ontologia, houve também uma análise dos dados dispostos pela Plataforma Nilo Peçanha e pelos repositórios de indicadores educacionais do INEP. Essa ação baseou-se no entendimento pelo pesquisador de que o enriquecimento dos dados institucionais a partir da sua integração com conteúdos externos teria capacidade de auxiliar o processo de construção de novas formas de compreensão das relações entre os construtos envolvidos na tese.

A atividade seguinte buscou especificar em linguagem natural as principais características da ontologia quanto aos seus componentes e relacionamentos, além de definir formalmente os conceitos identificados em seu escopo. Reuniões virtuais foram realizadas entre o pesquisador e as duas especialistas de domínio para discutir diversos aspectos acerca dos três construtos envolvidos na pesquisa e, assim, construir uma compreensão mútua de quais conceitos poderiam integrar a ontologia de maneira que sua estrutura estivesse coerente com as questões norteadoras do estudo.

A especificação da ontologia ficou assim definida:

Cada **Campus** do IFMG oferta pelo menos um **curso** de graduação, os quais certificam determinado **grau acadêmico** e pertencem a uma **área de conhecimento**. Os cursos também são caracterizados pelo turno no qual é ofertado, o último conceito no Enade, a relação aluno x professor e o percentual médio de estudantes cuja renda familiar *per capita* é de até um salário mínimo e meio. Um **estudante** é uma **pessoa** com cor e gênero que pode se matricular em um curso após aprovação em um processo seletivo conduzido de acordo com as diferentes **formas de ingresso** da instituição. A matrícula do estudante é caracterizada pela idade do mesmo no momento do ingresso, o ano de ingresso, a carga horária total da matriz curricular da turma, a diferença de tempo (em anos) entre o ingresso e o término do Ensino Médio e quantidade total de períodos cursados. A partir deste momento, a situação da matrícula indica que o estudante está **matriculado** regularmente no curso. Quando o estudante conclui todas as exigências acadêmicas do curso, a situação de matrícula indica que ele está **formado**. Quando o estudante interrompe seu curso, seja permanente ou temporariamente (como nos casos de trancamentos), a situação de matrícula indica que ele está **evadido**. Quando o estudante participa de processos seletivos da Assistência Estudantil para obtenção de auxílio, sua condição indica demanda e a solicitação é deferida, ele é identificado como **bolsista**. Estudantes bolsistas recebem um auxílio socioeconômico conhecido como bolsa permanência (BP1, BP2, BP3 ou BP4). Os estudantes são caracterizados, ainda, pelo indicativo se nasceu na cidade onde o curso é ofertado. O resultado da atuação do estudante durante a participação nas atividades

acadêmicas enquanto ele está matriculado em um curso promove o seu **desempenho acadêmico**, fornecendo informações sobre o coeficiente de rendimento geral, coeficiente de rendimento do primeiro semestre letivo, coeficiente de rendimento do primeiro ano letivo, percentual de reprovações durante todo o curso, percentual de reprovações no primeiro semestre letivo, percentual de reprovações no primeiro ano letivo, coeficiente de rendimento antes da pandemia, coeficiente de rendimento durante a pandemia e quantidade de disciplinas aprovadas e reprovadas.

O Quadro 4 apresenta a listagem dos conceitos envolvidos na ontologia, sua definição e a base teórica utilizada para uniformização do entendimento.

Quadro 4: Definição dos conceitos da ontologia

Conceito	Definição	Base teórica
Evasão escolar	Fenômeno complexo vinculado à desistência dos estudos por qualquer razão, cuja ocorrência é impactada por múltiplas variáveis que atuam interna ou externamente à instituição de ensino, fazendo com que as condições mínimas para a permanência do estudante sejam inviabilizadas.	Fritsch, Rocha e Vitelli (2015)
Auxílio socioeconômico	Benefício financeiro na forma de bolsa permanência concedida ao estudante em situação de vulnerabilidade socioeconômica com o objetivo de favorecer sua permanência na educação superior.	Manual de preenchimento do Censo da Educação Superior 2021 (aluno)
Desempenho acadêmico	Resultado de ações conduzidas por determinado estudante em seu curso e avaliadas pelos educadores em termos de rendimento e eficiência, sendo expresso no formato de coeficiente de rendimento.	Munhoz (2004) e Miranda et al. (2015)
Campus	Local de realização das atividades acadêmicas dos cursos presenciais ofertados pela instituição.	Manual de preenchimento do Censo da Educação Superior 2021 (IES)
Curso	Combinação de disciplinas e atividades organizadas em campos gerais ou específicos do conhecimento, para atender objetivos educacionais definidos pela instituição, em consonância com diretrizes curriculares.	Glossário de Legislação do Ensino Superior da UFPA
Grau acadêmico	Grau conferido a um curso de graduação, podendo ser: bacharelado, licenciatura ou	Manual de preenchimento do Censo da Educação Superior 2021 (curso)

	tecnológico.	
Área de conhecimento do curso	Classificação desenvolvida pela CAPES compreendendo oito grandes áreas do conhecimento, as quais os cursos estão vinculados.	Glossário de termos referentes ao ensino de graduação do CEFET-MG
Estudante	Indivíduo regularmente matriculado em um curso ou que já ingressou na instituição mas não está cursando nenhuma disciplina (como nos casos de estudantes formados ou evadidos).	Glossário de Legislação do Ensino Superior da UFPA
Estudante bolsista	Estudante matriculado que cumpriu os requisitos da Política de Assistência Estudantil e teve sua solicitação de auxílio socioeconômico deferida em processo seletivo institucional, tendo sido atribuído um tipo de bolsa permanência.	Política de Assistência Estudantil do IFMG e Glossário do painel de dados da UFRGS
Matrícula	Ato formal de vinculação de um estudante a um curso de graduação. Quando o vínculo está ativo em decorrência do ato da matrícula (inicial ou de renovação), o estudante está matriculado . Quando o vínculo não está ativo, o estudante pode estar formado (concluiu com sucesso todas as exigências acadêmicas) ou evadido (não houve continuidade da formação educacional, temporária ou definitiva).	Regulamento de Ensino dos Cursos de Graduação do IFMG, SEMERJ (2023) , Glossário de Legislação do Ensino Superior da UFPA, José, Broilo e Andreoli (2011) e Comissão Especial de Estudos Sobre a Evasão nas Universidades Públicas Brasileiras (1996)
Forma de ingresso	Tipo de processo seletivo ao qual o estudante se submeteu para ingressar em determinado curso.	Manual de preenchimento do Censo da Educação Superior 2021 (aluno)

Fonte: Elaborado pelo autor (2023)

De acordo com a SABiO, a demanda por modularizar ontologias precisa envolver a análise de três aspectos: independência, coesão e tamanho. Considerando a necessidade de proximidade entre os três construtos englobados na ontologia e a criação de um artefato com tamanho reduzido, facilitando assim a sua compreensão e utilização, entendeu-se que a modularização dos elementos não traria benefícios diretos que contribuíssem para o alcance dos objetivos propostos pela pesquisa.

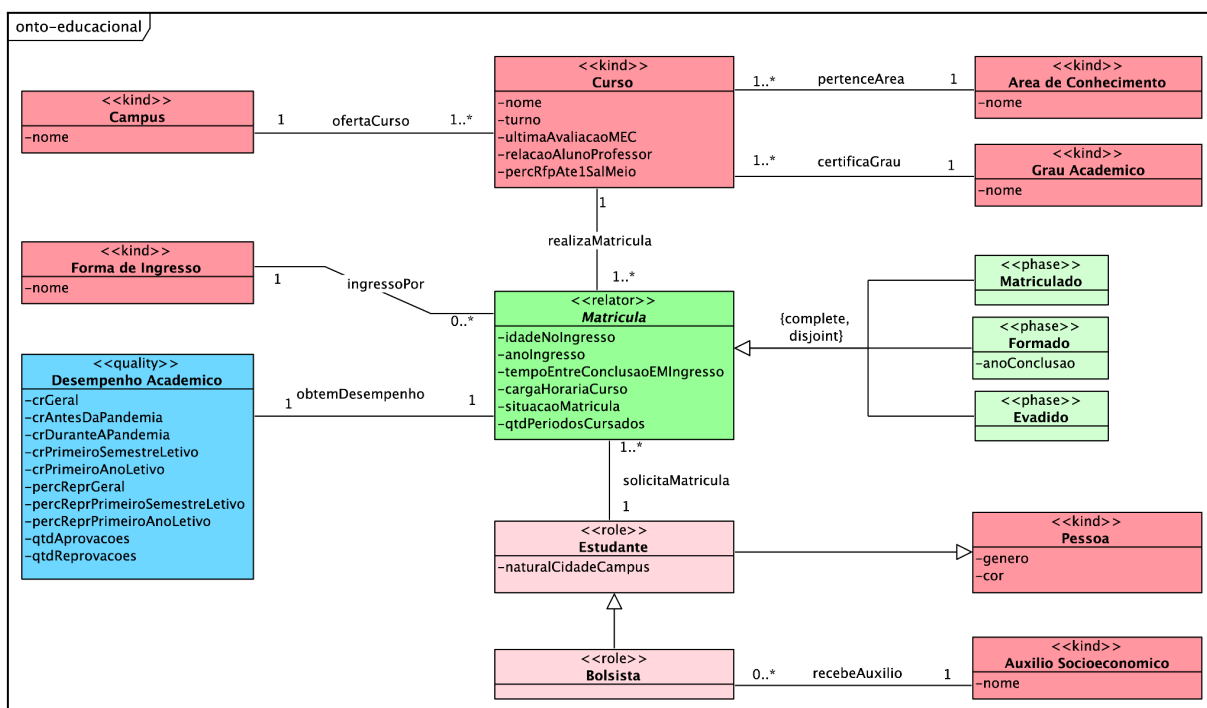
As buscas para identificar artefatos ontológicos que pudessem, eventualmente, ser reutilizados na tese, retornaram a ontologia [Curriculum Course Syllabus Ontology \(CCSO\)](#)⁹.

⁹ <https://vkreations.github.io/CCSO/>

Trata-se de uma ontologia concebida para contribuir com o entendimento sobre conceitos e entidades do ambiente acadêmico, notadamente aqueles relacionados com a organização administrativa de uma instituição educacional. Embora com maior abrangência comparativamente à ontologia desta tese, a CCSO permitiu compreender como ocorrem as relações entre os tipos Pessoa e Estudante em outro contexto, bem como a vinculação do Estudante a um Curso e a obtenção de seus resultados de aprendizagem. Outra diferença marcante refere-se à grande quantidade de propriedades que caracterizam cada tipo da CCSO, inclusive alguns reutilizados a partir de outros artefatos, como [FOAF](http://xmlns.com/foaf/0.1/)¹⁰ e [Schema.org](https://schema.org/)¹¹.

O encerramento da segunda fase ocorreu com a elaboração de um modelo baseado na linguagem OntoUML representando visualmente todos os conceitos especificados para a ontologia de referência e as relações entre eles. Tal construção foi realizada em várias iterações que, em determinadas ocasiões, contaram com a participação das especialistas de domínio no sentido de sugerir melhorias ou consensuar o entendimento sobre algum aspecto particular. A Figura 41 apresenta a versão final da ontologia de referência, denominada “Onto-Educacional”.

Figura 41: Ontologia de referência elaborada com OntoUML



Fonte: Elaborada pelo autor (2023)

¹⁰ <http://xmlns.com/foaf/0.1/>

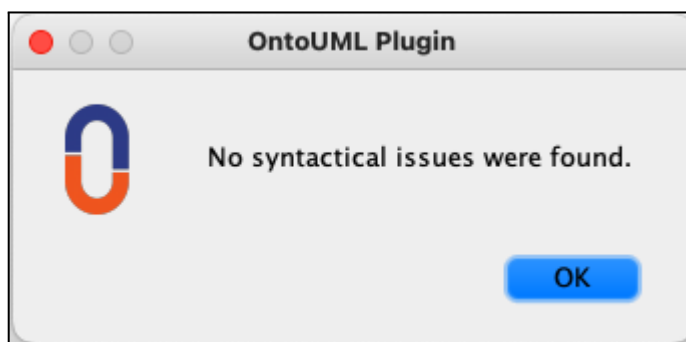
¹¹ <https://schema.org/>

É possível notar que a maior parte das classes que compõem a ontologia de referência foram estereotipadas como *kinds*, indicando que se referem a conceitos dotados do princípio uniforme de identidade. A classe *Curso* se relaciona diretamente com as classes *Campus* (para identificar o local de oferta), *Área de Conhecimento* (para caracterizar a área de estudos do curso) e *Grau Acadêmico* (para classificar o curso conforme a certificação final dos estudantes). A classe *Estudante* possui uma relação de especialização com a classe *Pessoa*, tendo recebido o estereótipo *role* devido ao fato deste conceito representar um papel executado pelo indivíduo dentro do contexto acadêmico em determinado momento e sob condições específicas.

O estereótipo *role* foi aplicado à classe *Bolsista*, pois se trata de um papel com aplicação dependente de outros fatores, como a aprovação em processo seletivo da Assistência Estudantil. Quando existe a instanciação de um bolsista, têm-se a definição de qual auxílio o mesmo receberá, simbolizado pela classe *Auxílio Socioeconômico*. A relação entre um estudante e seu curso é mediada pelo *relator Matrícula*, permitindo abstrair os dados próprios do relacionamento, bem como integrá-lo com a classe *Forma de Ingresso*. As mudanças de estado de uma matrícula são representadas por três classes disjuntas e completas estereotipadas como *phase*, um tipo anti-rígido, apontando se ela está ativa (*Matriculado*) ou encerrada (*Formado* ou *Evadido*). A abstração dos dados que representam o desempenho acadêmico de um estudante qualificam a sua matrícula em um curso, motivo pelo qual foi definida a classe *Desempenho Acadêmico*, estereotipada como *quality*.

A verificação de adequação do modelo às regras estipuladas pela UFO envolveu o uso de uma funcionalidade disponível no *plugin* da linguagem OntoUML para o software *Visual Paradigm*, onde é analisada a coerência sintática dos elementos e suas relações com um conjunto de regras e restrições formais. De acordo com [Fonseca et al. \(2021\)](#) e [Guizzardi et al. \(2021\)](#), quando alguma inconsistência é encontrada pelo processo de verificação, o *plugin* apresenta ao usuário detalhes sobre esse desvio de acordo com o seu nível de gravidade, erro (inconsistências ontológicas ou lógicas) ou aviso (problemas menos críticos, como a subespecificação de um modo, por exemplo). Após submeter a Onto-Educacional ao processo de verificação, nenhuma inconsistência foi encontrada, conforme observa-se na Figura 42.

Figura 42: Verificação sintática da ontologia conforme as regras da OntoUML



Fonte: Elaborada pelo autor (2023)

4.3.3 Design, implementação e testes

O escopo da terceira fase compreendeu atividades relacionadas ao design e à implementação da ontologia operacional, tendo como base a ontologia de referência construída na fase anterior, abrangendo ainda atividades para disponibilização do conjunto de dados em um ambiente eletrônico e de testes para avaliação da ontologia elaborada.

4.3.3.1 Design e implementação da ontologia operacional

O *plugin* da linguagem OntoUML existente no software *Visual Paradigm* contém uma funcionalidade para exportar o modelo criado nessa linguagem para um arquivo com as especificações da gUFO. Contudo, esta ação somente pode ser executada quando o modelo não apresenta nenhuma inconsistência sintática, como no caso da Onto-Educacional. Utilizando essa funcionalidade, foi gerado um arquivo no formato OWL e sintaxe *Turtle* contendo os elementos da ontologia especificada em OntoUML e as definições sobre reutilização de recursos da UFO em tais elementos, como *Kinds*, *Roles* e *Phases*. A Figura 43 contém um fragmento do arquivo, no qual é possível observar a especificação da classe “Pessoa” integrada com recursos da OWL e da gUFO.

Figura 43: Fragmento de código do arquivo com a ontologia exportada em gUFO

```

onto-educacional:Pessoa rdf:type owl:Class, gufo:Kind, owl:NamedIndividual;
rdfs:subClassOf gufo:FunctionalComplex;
rdfs:label "Pessoa"@en.

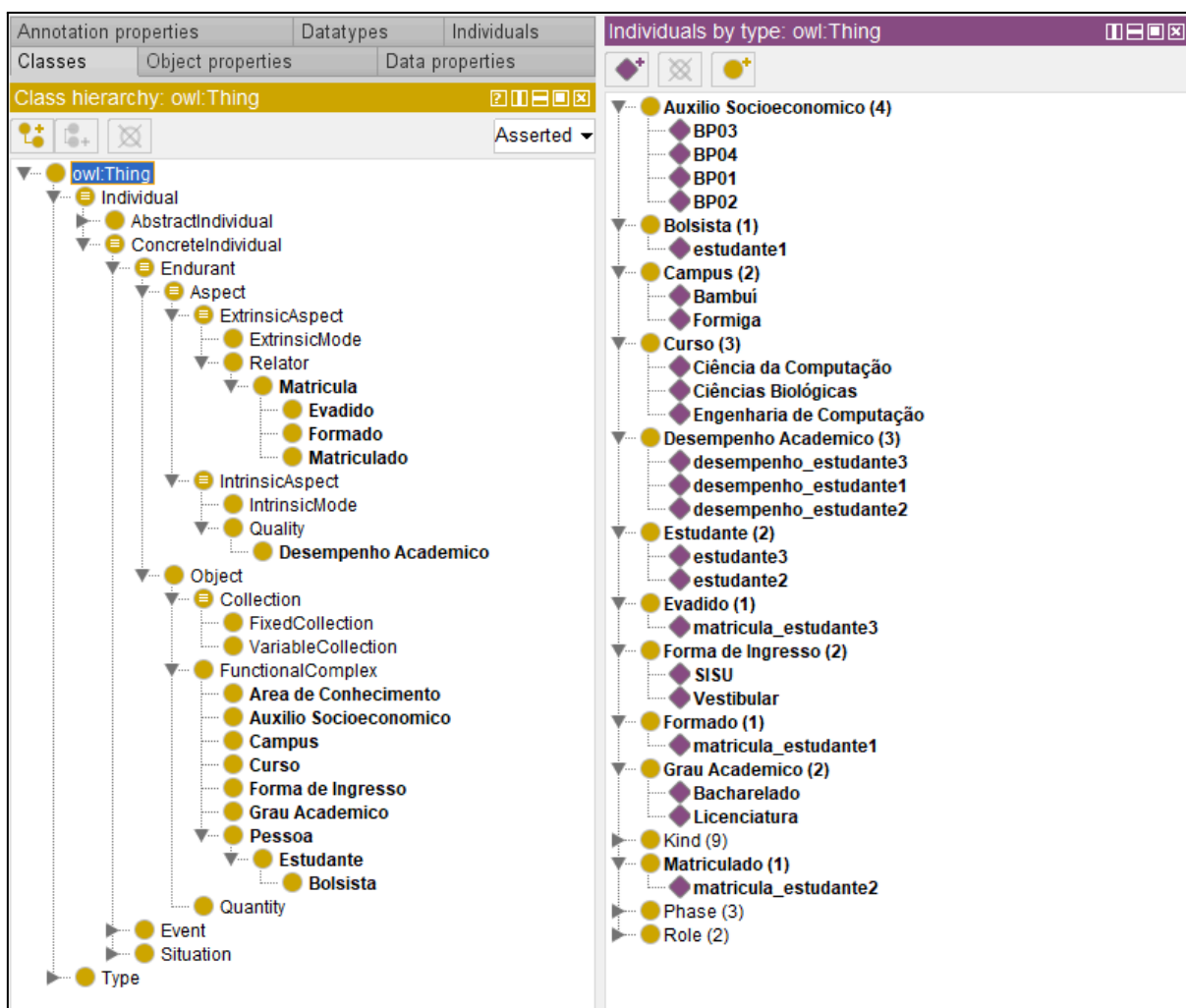
```

Fonte: Elaborada pelo autor (2023)

A continuidade da transformação da ontologia de referência em ontologia operacional demandou a criação de um novo projeto no software *Protégé*, tendo como base o arquivo exportado pelo *Visual Paradigm*. É importante mencionar que foi necessário importar, manualmente, a gUFO no *Protégé* para que as classes da Onto-Educacional tivessem a

devida adequação na taxonomia de indivíduos. Na Figura 44 verifica-se a integração entre as classes próprias da Onto-Educacional e aquelas que compõem a gUFO (quadro “Class hierarchy”), assim como a instanciação das classes de tipos (quadro “Individuals by type”).

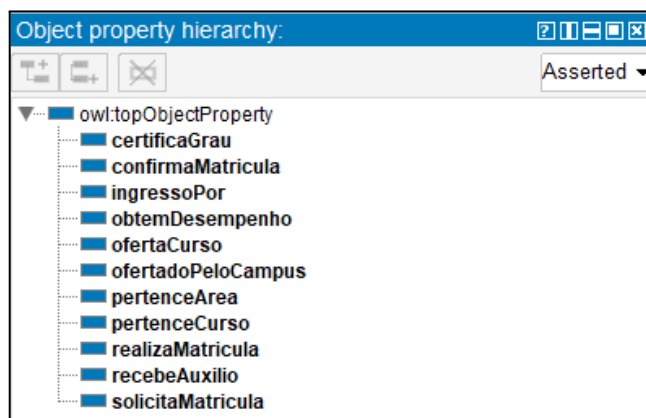
Figura 44: Integração das classes da Onto-Educacional com a gUFO



Fonte: Elaborada pelo autor (2023)

As propriedades dos objetos são apresentadas na Figura 45, permitindo caracterizar os diversos relacionamentos entre os conceitos da ontologia. Com o intuito de facilitar futuros processos de recuperação da informação, optou-se pela criação de três propriedades inversas: *pertenceCurso* (inversa de *realizaMatricula*), *confirmaMatricula* (inversa de *solicitaMatricula*) e *ofertadoPeloCampus* (inversa de *ofertaCurso*).

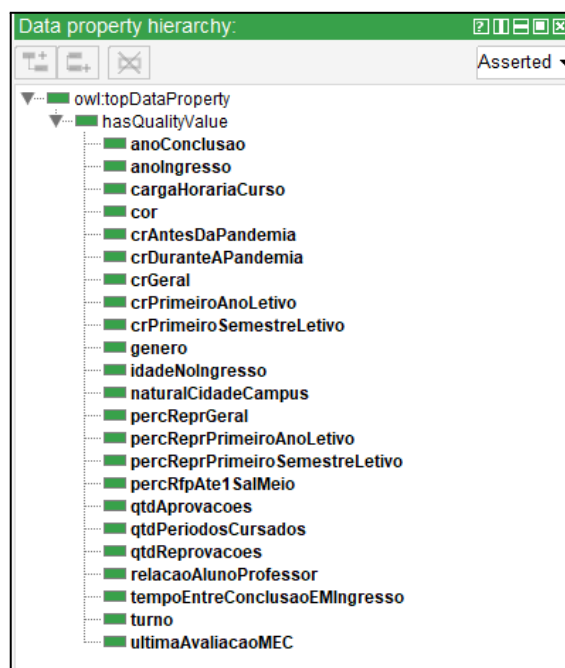
Figura 45: Propriedades dos objetos da ontologia



Fonte: Elaborada pelo autor (2023)

Observa-se na Figura 46 as propriedades dos dados, cuja definição vem ao encontro da especificação da ontologia construída na fase de captura e formalização da ontologia (Subseção 4.3.2). Tais dados possuem importância fundamental para que as questões de competência da pesquisa sejam respondidas e, assim, permitam avaliar a adequação do artefato ontológico enquanto suporte para integração de dados oriundos de fontes heterogêneas.

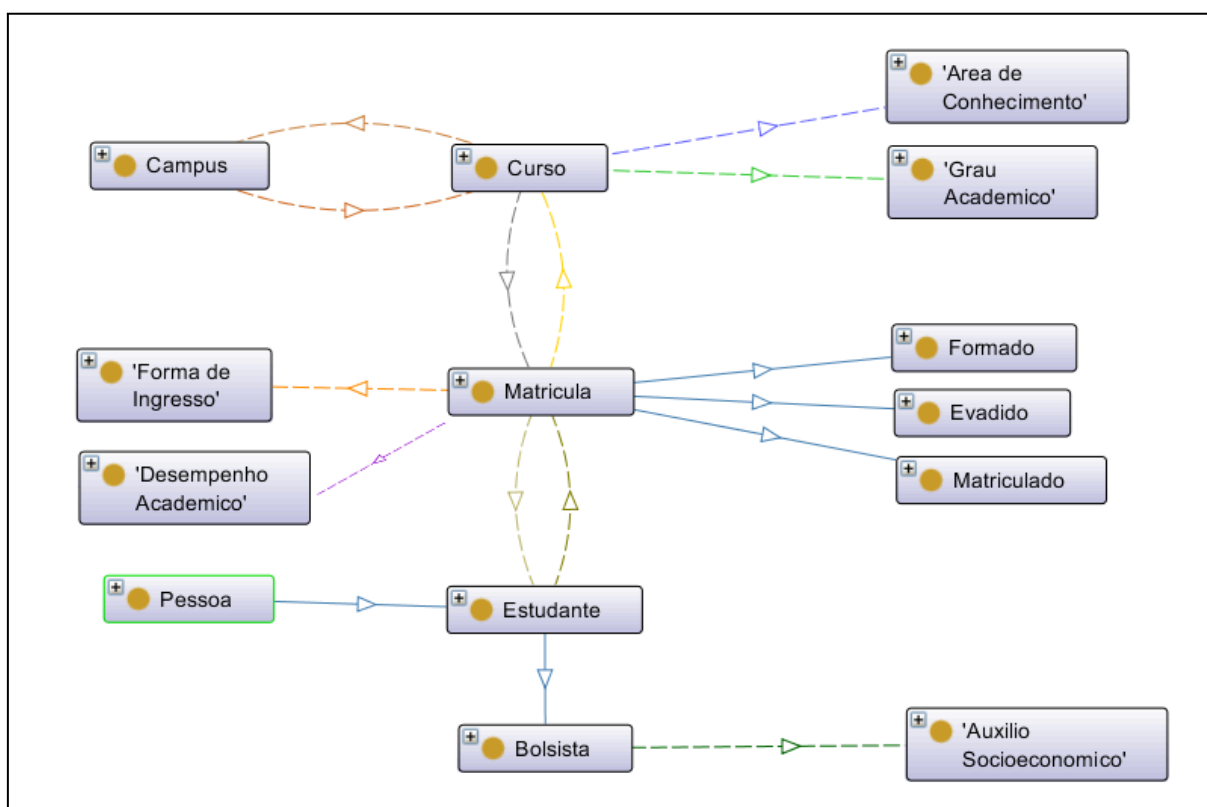
Figura 46: Propriedades dos dados da ontologia



Fonte: Elaborada pelo autor (2023)

A ontologia operacional elaborada no *Protégé* está representada graficamente na Figura 47, destacando as classes que representam os conceitos advindos dos construtos abordados na pesquisa e que possibilitam organizar as informações sobre o domínio em questão. As relações entre as classes também são representadas para indicar como está definida a vinculação de um conceito a outro na ontologia. Nenhum novo conceito foi adicionado em comparação com a ontologia de referência.

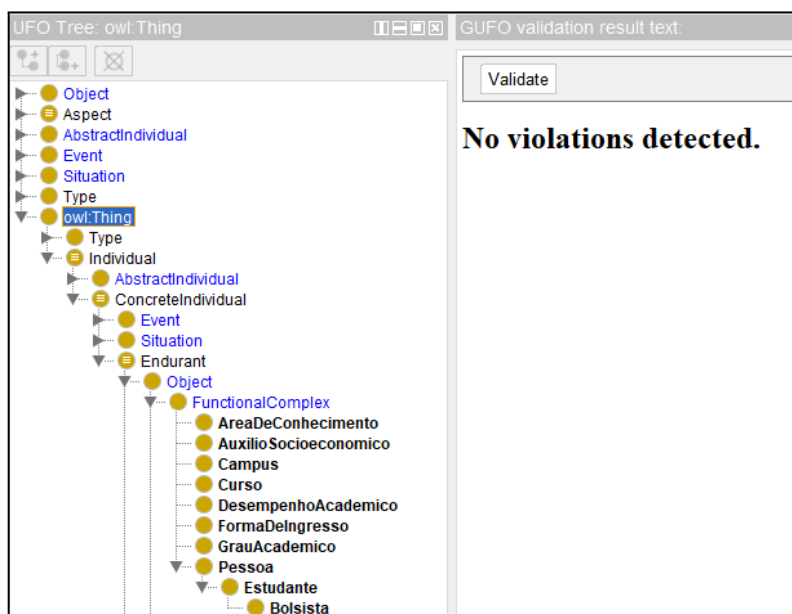
Figura 47: Representação gráfica da ontologia operacional



Fonte: Elaborada pelo autor (2023)

O *plugin* da gUFO para o Protégé permite que a ontologia operacional seja verificada no que concerne às regras sintáticas da UFO, procedimento semelhante ao conduzido anteriormente com a ontologia de referência. Observa-se na Figura 48 que nenhuma inconsistência foi encontrada na Onto-Educacional, estando a mesma, portanto, de acordo com os princípios metafísicos da UFO.

Figura 48: Verificação sintática da ontologia operacional conforme a UFO



Fonte: Elaborada pelo autor (2023)

Encerradas as atividades de design e implementação da ontologia operacional, a versão final foi publicada em OWL no repositório virtual do projeto: <https://github.com/educardosomelo/onto-educacional/tree/main/Ontologia>.

4.3.3.2 Coleta e construção do conjunto de dados

A ontologia operacional elaborada na atividade anterior forneceu a estruturação necessária para que ficasse evidente o conjunto de dados a ser obtido nos ambientes informacionais do IFMG, da Plataforma Nilo Peçanha e do INEP. As buscas consideraram o escopo definido para esta tese: turmas de ingressantes entre o primeiro semestre letivo de 2018 e o segundo semestre letivo de 2022 de cursos superiores presenciais ofertados pelos *campi* do IFMG. Na Subseção 4.3.2 foi mencionado que o IFMG disponibilizou arquivos com seus dados acadêmicos delimitados com aderência a este escopo: uma planilha eletrônica com atributos caracterizadores dos estudantes e um arquivo CSV contendo registros das disciplinas cursadas por eles. Tais arquivos foram utilizados como fonte dos dados do IFMG para a condução da pesquisa.

Em seguida, procedeu-se com a criação da estrutura do repositório intermediário dos dados no SGBD *MySQL*, proporcionando um ambiente para tratamento e integração dos dados brutos. Foi criada uma tabela chamada “aluno” que recebeu, inicialmente, os dados acadêmicos existentes na planilha eletrônica disponibilizada pelo IFMG. A tabela possuía os seguintes atributos: *identificador do estudante*, *campus*, *curso*, *data de nascimento*, *gênero*, *naturalidade*, *cor*, *forma de ingresso*, *ano de encerramento do Ensino Médio*, *ano de*

ingresso no curso, ano de conclusão do curso, turno de estudos, carga horária total do curso, coeficiente de rendimento geral e situação atual da matrícula. A criação de outra tabela, chamada “disciplinas”, serviu para armazenar os dados advindos do arquivo CSV, representando os registros das disciplinas cursadas pelos estudantes. Sua estrutura contempla os atributos: *identificador do estudante, campus, curso, período letivo, código da disciplina, nome da disciplina, nota final e situação final na disciplina.*

Com os dados acadêmicos do IFMG disponíveis no repositório intermediário, tarefas de pré-processamento foram executadas para que a sua qualidade sofresse incremento, reduzindo assim a possibilidade de erros interpretativos futuros quando do estudo sobre as relações entre os três construtos analisados nesta tese. As atualizações de registros e criação de novos atributos aconteceram na tabela “aluno”. Registros duplicados não foram encontrados nas duas tabelas; para alguns estudantes, havia mais de um registro de vínculo com uma mesma disciplina, porém com situação de matrícula diferente, indicando que foi necessário cursá-la mais de uma vez tentando a aprovação. Nenhum dado faltante foi identificado nos registros dos alunos e das disciplinas, possivelmente em função do preenchimento dos campos no sistema acadêmico do IFMG ser obrigatório. No que se refere à padronização de valores em cada atributo, alguns passaram por ajustes. Registros com o atributo *situação atual da matrícula* igual a “Evasão”, “Transferência” ou “Desligado” foram alterados para “Evadido”, enquanto aqueles com valor igual “Concluído” passaram a ter o valor “Formado”. O atributo *turno de estudos* teve registros alterados quando o valor era igual a “Manhã/Tarde”, trocando para “Integral”. Os registros que possuíam o valor “Processo seletivo interno” no atributo *forma de ingresso* foram mudados para “Vestibular” após consulta realizada junto à Pró-Reitoria de Ensino do IFMG.

Alguns atributos originais passaram por um processo de transformação com a construção de novos atributos contendo informações mais relevantes para a pesquisa, sendo os originais excluídos ao final dessas tarefas. O atributo *naturalidade* serviu de base para criação do atributo *natural cidade campus*, indicando se o estudante nasceu na mesma cidade do Campus com o qual estava vinculado. O atributo *data de nascimento* foi transformado em um novo atributo denominado *idade no ingresso* que representa a idade do estudante no momento de início do seu vínculo com algum curso (matrícula). A subtração dos valores encontrados nos atributos *ano de ingresso no curso* e *ano de encerramento do Ensino Médio* originou o atributo *tempo entre conclusão do Ensino Médio e ingresso no curso*.

Visando atender as relações existentes na ontologia operacional entre a classe Curso e as classes Área de Conhecimento e Grau Acadêmico, dois novos atributos foram criados: *grau* (com valores possíveis entre “Bacharelado”, “Licenciatura” e “Tecnológico”) e *área de conhecimento* (de acordo com a [Tabela de Áreas de Conhecimento/Avaliação da](#)

[CAPES](#)). Esses novos atributos foram atualizados via *script* nos registros da tabela “aluno” de acordo com o valor do atributo *curso*.

Com o apoio de *scripts* elaborados em *Python*, os registros de disciplinas cursadas (localizados na tabela “disciplinas”) foram manipulados para a criação e preenchimento de atributos indicadores do desempenho acadêmico dos estudantes na tabela “aluno”. O primeiro deles, denominado *quantidade de períodos cursados*, representa o número de períodos letivos cursados pelo estudante, independentemente se ele foi finalizado por completo. Outros dois atributos, nomeados como *quantidade de aprovações* e *quantidade de reprovações*, armazenam, respectivamente, o número de disciplinas com status de aprovação e reprovação do estudante. Utilizando a nota final de cada disciplina cursada, foi calculado para cada estudante os valores dos atributos *coeficiente de rendimento antes da pandemia* (para disciplinas cursadas em 2018 e 2019), *coeficiente de rendimento durante a pandemia* (para disciplinas cursadas em 2020, 2021 e 2022), *coeficiente de rendimento do primeiro semestre letivo* (indicando qual foi o rendimento acadêmico médio apenas nas disciplinas realizadas durante o primeiro semestre letivo do curso) e *coeficiente de rendimento do primeiro ano letivo* (similar ao atributo anterior, porém considerando disciplinas cursadas ao longo do primeiro ano letivo). Com base na situação final de cada disciplina cursada pelos estudantes, os seguintes atributos foram criados e calculados: *percentual de reprovações geral* (tendo em conta todas as disciplinas realizadas pelo estudante em determinado curso), *percentual de reprovações no primeiro semestre letivo* e *percentual de reprovações no primeiro ano letivo* (ambas considerando a quantidade de disciplinas reprovadas em função do total de disciplinas cursadas no semestre ou no ano letivo).

Para promover a integração dos dados acadêmicos já armazenados no repositório do *MySQL* com os dados da assistência estudantil, uma consulta foi realizada no sistema informatizado de processos seletivos desse setor para obter o identificador de todos aqueles estudantes que foram contemplados com auxílio socioeconômico no formato de bolsa permanência desde 2018 até o final de 2022, além do tipo de bolsa atribuída para cada estudante. Esses dados foram inseridos na tabela “aluno” em dois novos atributos denominados *bolsista da assistência estudantil* e *auxílio socioeconômico recebido*, tendo o atributo *identificador do estudante* atuado como elemento integrador dos dados existentes aos obtidos no sistema da assistência estudantil. Importa ressaltar que, encerrado o procedimento de integração, os valores que identificavam os estudantes do conjunto de dados foram substituídos por números aleatórios, promovendo a anonimização dos registros.

A última tarefa de pré-processamento dos dados consistiu no enriquecimento dos registros a partir de sua integração com informações advindas da Plataforma Nilo Peçanha e

da central de dados abertos do INEP. Três atributos foram criados e tiveram seus valores inseridos via *script* nos registros da tabela “aluno”: *última avaliação do MEC* (último conceito do curso no ENADE ou na avaliação presencial), *relação aluno professor* (índice que permite avaliar a eficiência da oferta do curso) e *percentual de alunos do curso com renda per capita familiar de até um salário mínimo e meio* (utilizando a mesma base de corte concebida como critério do PNAES).

O conjunto de dados mapeado conforme a ontologia operacional ficou com 10.263 registros, compostos de 32 atributos (11 originais e 21 criados). O seu detalhamento é apresentado no Quadro 5.

Quadro 5 - Atributos mapeados conforme a ontologia operacional

Nº	Atributo	Tipo	Criado	Descrição/valores possíveis	Mapeamento com a ontologia operacional
1	id	Numérico	Sim	Código sequencial identificador do estudante.	<i>onto-educacional:Estudante</i> <i>rdf:type owl:Class</i> ; <i>rdfs:subClassOf onto-educacional:Pessoa</i> .
2	campus	Nominal	Não	Nome do Campus de vínculo do estudante.	<i>onto-educacional:Campus</i> <i>rdf:type owl:Class</i> ; <i>rdfs:subClassOf gufo:FunctionalComplex</i> .
3	curso	Nominal	Não	Nome do curso no qual o estudante se matriculou.	<i>onto-educacional:Curso</i> <i>rdf:type owl:Class</i> ; <i>rdfs:subClassOf gufo:FunctionalComplex</i> .
4	grau_academico	Catagórico	Sim	Grau acadêmico conferido pelo curso. <i>Bacharelado</i> <i>Licenciatura</i> <i>Tecnológico</i>	<i>onto-educacional:GrauAcademico</i> <i>rdf:type owl:Class</i> ; <i>rdfs:subClassOf gufo:FunctionalComplex</i> .
5	area_conhecimento	Catagórico	Sim	Área do curso de acordo com a CAPES. <i>Ciências Agrárias</i> <i>Ciências Biológicas</i> <i>Ciências Exatas e da Terra</i> <i>Ciências Humanas</i> <i>Ciências Sociais Aplicadas</i> <i>Engenharias</i> <i>Linguística, Letras e Artes</i> <i>Ciências da Saúde</i>	<i>onto-educacional:AreaDeConhecimento</i> <i>rdf:type owl:Class</i> ; <i>rdfs:subClassOf gufo:FunctionalComplex</i> .
6	turno	Catagórico	Não	Turno de oferta do curso no Campus. <i>Integral</i> <i>Matutino</i> <i>Noturno</i> <i>Vespertino</i>	<i><curso:turno></i> <i>rdf:type owl:DatatypeProperty</i> ; <i>rdfs:subPropertyOf gufo:hasQualityValue</i> ; <i>rdfs:domain onto-educacional:Curso</i> ; <i>rdfs:label "turno"@pt</i> .
7	carga_horaria_curso	Numérico	Não	Carga horária total do curso (conforme matriz curricular).	<i><matricula:cargaHorariaCurso></i> <i>rdf:type owl:DatatypeProperty</i> ; <i>rdfs:subPropertyOf gufo:hasQualityValue</i> ; <i>rdfs:domain onto-educacional:Matricula</i> ; <i>rdfs:label "cargaHorariaCurso"@pt</i> .
8	ultima_avaliacao_mec	Numérico	Sim	Último conceito do curso no ENADE ou na avaliação presencial feita pelo MEC.	<i><curso:ultimaAvaliacaoMEC></i> <i>rdf:type owl:DatatypeProperty</i> ; <i>rdfs:subPropertyOf gufo:hasQualityValue</i> ; <i>rdfs:domain onto-educacional:Curso</i> ; <i>rdfs:label "ultimaAvaliacaoMEC"@pt</i> .

9	relacao_aluno_professor	Numérico	Sim	Proporção entre os estudantes matriculados e os professores de um curso.	<curso:relacaoAlunoProfessor> rdfs:type owl:DatatypeProperty ; rdfs:subPropertyOf gufo:hasQualityValue ; rdfs:domain onto-educacional:Curso ; rdfs:label "relacaoAlunoProfessor"@pt .
10	perc_rfp_ate_1salmeio_curso	Numérico	Sim	Percentual de alunos do curso com renda <i>per capita</i> familiar de até um salário mínimo e meio.	<curso:percRfpAte1SalMeio> rdfs:type owl:DatatypeProperty ; rdfs:subPropertyOf gufo:hasQualityValue ; rdfs:domain onto-educacional:Curso ; rdfs:label "percRfpAte1SalMeio"@pt .
11	forma_ingresso	Catagórico	Não	Forma de ingresso do estudante no curso. <i>Obtenção de novo título</i> <i>SISU</i> <i>Transferência externa</i> <i>Vestibular</i>	onto-educacional:FormaDeIngresso rdfs:type owl:Class ; rdfs:subClassOf gufo:FunctionalComplex .
12	genero	Catagórico	Não	Gênero do estudante. <i>F (Feminino)</i> <i>M (Masculino)</i>	<peessoa:genero> rdfs:type owl:DatatypeProperty ; rdfs:subPropertyOf gufo:hasQualityValue ; rdfs:domain onto-educacional:Pessoa ; rdfs:label "genero"@pt .
13	cor	Catagórico	Não	Cor do estudante. <i>Amarela</i> <i>Branca</i> <i>Indígena</i> <i>Parda</i> <i>Preta</i>	<peessoa:cor> rdfs:type owl:DatatypeProperty ; rdfs:subPropertyOf gufo:hasQualityValue ; rdfs:domain onto-educacional:Pessoa ; rdfs:label "cor"@pt .
14	natural_cidade_campus	Catagórico	Sim	Indica se o estudante nasceu na mesma cidade do Campus onde o seu curso é ofertado. <i>N (Não)</i> <i>S (Sim)</i>	<estudante:naturalCidadeCampus> rdfs:type owl:DatatypeProperty ; rdfs:subPropertyOf gufo:hasQualityValue ; rdfs:domain onto-educacional:Estudante ; rdfs:label "naturalCidadeCampus"@pt .
15	idade_ingresso	Numérico	Sim	Idade do estudante no momento da matrícula no curso.	<matricula:idadeNoIngresso> rdfs:type owl:DatatypeProperty ; rdfs:subPropertyOf gufo:hasQualityValue ; rdfs:domain onto-educacional:Matricula ; rdfs:label "idadeNoIngresso"@pt .
16	ano_ingresso	Numérico	Não	Ano de ingresso do estudante no curso.	<matricula:anoIngresso> rdfs:type owl:DatatypeProperty ; rdfs:subPropertyOf gufo:hasQualityValue ; rdfs:domain onto-educacional:Matricula ; rdfs:label "anoIngresso"@pt .
17	tempo_entre_conclusao_em_ingresso	Numérico	Sim	Tempo (em anos) decorrido entre a conclusão do Ensino Médio pelo estudante e o seu ingresso no	<matricula:tempoEntreConclusaoEMIngresso> rdfs:type owl:DatatypeProperty ;

				curso.	<i>rdfs:subPropertyOf gufo:hasQualityValue ; rdfs:domain onto-educacional:Matricula ; rdfs:label "tempoEntreConclusaoEMIngresso"@pt .</i>
18	ano_conclusao	Numérico	Não	Ano de conclusão do curso (valor preenchido apenas para estudantes formados).	<i><formado:anoConclusao> rdfs:type owl:DatatypeProperty ; rdfs:subPropertyOf gufo:hasQualityValue ; rdfs:domain onto-educacional:Formado ; rdfs:label "anoConclusao"@pt .</i>
19	cr_geral	Numérico	Não	Coefficiente de rendimento geral do estudante no curso.	<i><desempenhoAcademico:crGeral> rdfs:type owl:DatatypeProperty ; rdfs:subPropertyOf gufo:hasQualityValue ; rdfs:domain onto-educacional:DesempenhoAcademico ; rdfs:label "crGeral"@pt .</i>
20	cr_antes_pandemia	Numérico	Sim	Coefficiente de rendimento do estudante no curso considerando apenas aquelas disciplinas cursadas em 2018 e 2019.	<i><desempenhoAcademico:crAntesDaPandemia> rdfs:type owl:DatatypeProperty ; rdfs:subPropertyOf gufo:hasQualityValue ; rdfs:domain onto-educacional:DesempenhoAcademico ; rdfs:label "crAntesDaPandemia"@pt .</i>
21	cr_durante_pandemia	Numérico	Sim	Coefficiente de rendimento do estudante no curso considerando apenas aquelas disciplinas cursadas em 2020, 2021 e 2022.	<i><desempenhoAcademico:crDuranteAPandemia> rdfs:type owl:DatatypeProperty ; rdfs:subPropertyOf gufo:hasQualityValue ; rdfs:domain onto-educacional:DesempenhoAcademico ; rdfs:label "crDuranteAPandemia"@pt .</i>
22	cr_primeiro_semestre_letivo	Numérico	Sim	Coefficiente de rendimento do estudante apenas nas disciplinas realizadas durante o primeiro semestre letivo do curso.	<i><desempenhoAcademico:crPrimeiroSemestreLetivo> rdfs:type owl:DatatypeProperty ; rdfs:subPropertyOf gufo:hasQualityValue ; rdfs:domain onto-educacional:DesempenhoAcademico ; rdfs:label "crPrimeiroSemestreLetivo"@pt .</i>
23	cr_primeiro_ano_letivo	Numérico	Sim	Coefficiente de rendimento do estudante apenas nas disciplinas realizadas durante o primeiro ano letivo do curso.	<i><desempenhoAcademico:crPrimeiroAnoLetivo> rdfs:type owl:DatatypeProperty ; rdfs:subPropertyOf gufo:hasQualityValue ; rdfs:domain onto-educacional:DesempenhoAcademico ; rdfs:label "crPrimeiroAnoLetivo"@pt .</i>
24	perc_repr_geral	Numérico	Sim	Percentual geral de reprovações (considerando todas as disciplinas realizadas pelo estudante em determinado curso).	<i><desempenhoAcademico:percReprGeral> rdfs:type owl:DatatypeProperty ; rdfs:subPropertyOf gufo:hasQualityValue ; rdfs:domain onto-educacional:DesempenhoAcademico ; rdfs:label "percReprGeral"@pt .</i>
25	perc_repr_primeiro_semestre_letivo	Numérico	Sim	Percentual de reprovações em disciplinas realizadas durante o primeiro semestre letivo do curso.	<i><desempenhoAcademico:percReprPrimeiroSemestreLetivo> rdfs:type owl:DatatypeProperty ;</i>

					<i>rdfs:subPropertyOf gufo:hasQualityValue ; rdfs:domain onto-educacional:DesempenhoAcademico ; rdfs:label "percReprPrimeiroSemestreLetivo"@pt .</i>
26	perc_repr_primeiro_ano_letivo	Numérico	Sim	Percentual de reprovações em disciplinas realizadas durante o primeiro ano letivo do curso.	<i><desempenhoAcademico:percReprPrimeiroAnoLetivo> rdfs:type owl:DatatypeProperty ; rdfs:subPropertyOf gufo:hasQualityValue ; rdfs:domain onto-educacional:DesempenhoAcademico ; rdfs:label "percReprPrimeiroAnoLetivo"@pt .</i>
27	qtd_periodos_cursados	Numérico	Sim	Quantidade de períodos letivos cursados pelo estudante no curso, inclusive aqueles não encerrados por completo.	<i><matricula:qtdPeriodosCursados> rdfs:type owl:DatatypeProperty ; rdfs:subPropertyOf gufo:hasQualityValue ; rdfs:domain onto-educacional:Matricula ; rdfs:label "qtdPeriodosCursados"@pt .</i>
28	qtd_aprovacoes	Numérico	Sim	Número de disciplinas realizadas com status de aprovação pelo estudante no curso.	<i><desempenhoAcademico:qtdAprovacoes> rdfs:type owl:DatatypeProperty ; rdfs:subPropertyOf gufo:hasQualityValue ; rdfs:domain onto-educacional:DesempenhoAcademico ; rdfs:label "qtdAprovacoes"@pt .</i>
29	qtd_reprovacoes	Numérico	Sim	Número de disciplinas realizadas com status de reprovação pelo estudante no curso.	<i><desempenhoAcademico:qtdReprovacoes> rdfs:type owl:DatatypeProperty ; rdfs:subPropertyOf gufo:hasQualityValue ; rdfs:domain onto-educacional:DesempenhoAcademico ; rdfs:label "qtdReprovacoes"@pt .</i>
30	bolsista_ae	Categórico	Sim	Indica se o estudante já foi contemplado com alguma bolsa de auxílio socioeconômico. N (Não) S (Sim)	<i>onto-educacional:Bolsista rdfs:type owl:Class ; rdfs:subClassOf onto-educacional:Estudante .</i>
31	auxilio_socioeconomico	Categórico	Sim	Tipo da bolsa concedida ao estudante selecionado pela Assistência Estudantil. BP1 (Bolsa permanência 1) BP2 (Bolsa permanência 2) BP3 (Bolsa permanência 3) BP4 (Bolsa permanência 4)	<i>onto-educacional:AuxilioSocioeconomico rdfs:type owl:Class ; rdfs:subClassOf gufo:FunctionalComplex .</i>
32	situacao_matricula	Categórico	Não	Indica a situação da matrícula do estudante em um curso. Evadido Formado Matriculado	<i>onto-educacional:Evadido rdfs:type owl:Class ; rdfs:subClassOf onto-educacional:Matricula . onto-educacional:Formado rdfs:type owl:Class ; rdfs:subClassOf onto-educacional:Matricula . onto-educacional:Matriculado rdfs:type owl:Class ; rdfs:subClassOf onto-educacional:Matricula .</i>

Fonte: Elaborado pelo autor (2023)

Embora o conjunto de dados construído no repositório intermediário tenha atributos e registros suficientes para serem explorados com o intuito de gerar informações, optou-se por caracterizá-lo de forma mais abrangente na Subseção 4.3.3.4 (Avaliação da ontologia), respondendo às questões de competência, e na Etapa 3 (Seção 4.4), momento em que dados suportaram as análises sobre as relações envolvendo os construtos da pesquisa.

4.3.3.3 Ingestão dos dados em um *Graph Database*

Encerradas as manipulações no conjunto de dados do repositório intermediário, iniciou-se a atividade para ingerir os registros em um banco de dados no formato de grafo. Após estudos sobre a utilização do *Neo4j* como plataforma, especialmente no que tange ao formato requerido para armazenamento dos dados em seu ambiente e sobre a linguagem padrão de execução de comandos (*Cypher*), ficou definido que as instâncias de algumas classes seriam inseridas individualmente via comandos, dada a quantidade pequena de registros e a não variação de seus valores no escopo do conjunto de dados, enquanto a inserção das instâncias de classes com muitos registros ocorreria por meio da importação de arquivos CSV. O repositório em grafo criado no *Neo4j* foi nomeado como *onto-educacional* e os *scripts* utilizados nesta atividade estão disponíveis no repositório virtual do projeto em <https://github.com/educardosomelo/onto-educacional/tree/main/Scripts>. Os arquivos CSV gerados para importação no *Neo4j* podem ser acessados em <https://github.com/educardosomelo/onto-educacional/tree/main/Dados>.

O comando *CREATE* permite inserir uma nova instância no repositório do *Neo4j*, informando inclusive suas possíveis propriedades. As classes, relações e propriedades criadas refletem a modelagem ontológica concluída anteriormente. As instâncias das classes *Campus*, *GrauAcademico*, *AreaDeConhecimento*, *FormaDeIngresso* e *AuxilioSocioeconomico* foram inseridas conforme o comando a seguir, no qual observa-se a indicação da nova instância (*cam1*), do nome da classe (*Campus*) e de uma propriedade (chamada *nome* e com valor *Arcos*).

```
CREATE (cam1:Campus {nome: 'Arcos'});
```

A classe *Curso* foi a primeira cujas instâncias tiveram origem em um arquivo CSV. De acordo com a ontologia operacional, um curso possui cinco propriedades específicas: *nome*, *turno*, *ultimaAvaliacaoMEC*, *relacaoAlunoProfessor* e *percRfpAte1SalMeio*. Para efeitos de identificação da instância, um código numérico (*idEstudante*) também compõe as propriedades dos cursos, sendo ele utilizado no relacionamento com a matrícula do estudante. O comando executado para a importação dos dados, exposto a seguir, cria uma nova instância para cada registro lido no arquivo CSV, preenchendo as propriedades com os dados advindos dos atributos do arquivo.

```

LOAD CSV WITH HEADERS FROM 'file:///Curso.csv' AS l
CREATE (cur:Curso {idEstudante: toInteger(l.id), nome:l.curso, turno:
l.turno, ultimaAvaliacaoMEC: toInteger(l.ultima_avaliacao_mec),
relacaoAlunoProfessor: toFloat(l.relacao_aluno_professor),
percRfpAte1SalMeio: toFloat(l.perc_rfp_ate_1salmeio_curso)})
RETURN COUNT(cur)

```

Considerando que cada instância possui seus próprios relacionamentos com outras instâncias, utilizou-se os comandos *MATCH* e *MERGE* para promover a ligação entre dois nós existentes no repositório. As instâncias de *Curso*, por exemplo, se relacionam com instâncias de *AreaDeConhecimento* para indicar que aquele curso pertence a determinada área do conhecimento. A execução da rotina de vinculação de instâncias no *Neo4j* requereu o carregamento do arquivo CSV com os dados dos cursos, no qual estava armazenada a área de conhecimento de cada um deles. O comando *MATCH* faz a busca no repositório por determinada instância, sendo que nesse caso ocorreram as buscas por um curso e por uma área de conhecimento, ambas com utilização do nome como critério de pesquisa. O comando *MERGE* promove, ao final de cada registro lido no arquivo CSV, a vinculação entre as duas instâncias encontradas, inclusive nomeando-a de acordo com a indicação feita no código. Observa-se, no comando a seguir, que as instâncias de *Curso* (chamada *cur*) e de *AreaDeConhecimento* (chamada *are*) foram vinculadas em uma relação nomeada como *pertenceArea*.

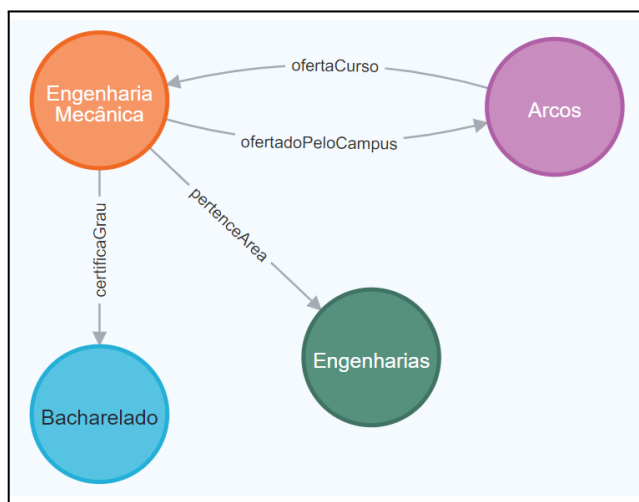
```

LOAD CSV WITH HEADERS FROM 'file:///Curso.csv' AS l
MATCH (cur:Curso {idEstudante: toInteger(l.id)})
MATCH (are:AreaDeConhecimento {nome: l.area_conhecimento})
MERGE (cur)-[:pertenceArea]->(are)

```

As instâncias de *Curso* ainda foram vinculadas com instâncias das classes *GrauAcademico* (relação *certificaGrau*) e *Campus* (relação direta *ofertadoPeloCampus* e inversa *ofertaCurso*). A Figura 49 apresenta, no formato de grafo, uma consulta executada no *Neo4j Browser* contendo a instância específica de um curso e suas relações com outras instâncias. Observa-se, de forma visual ao invés de apenas textual (tal como era possível quando os dados estavam armazenados no repositório intermediário em *MySQL*), que o curso de Engenharia Mecânica certifica o grau de Bacharelado, pertence à área das Engenharias e é ofertado pelo Campus localizado na cidade de Arcos.

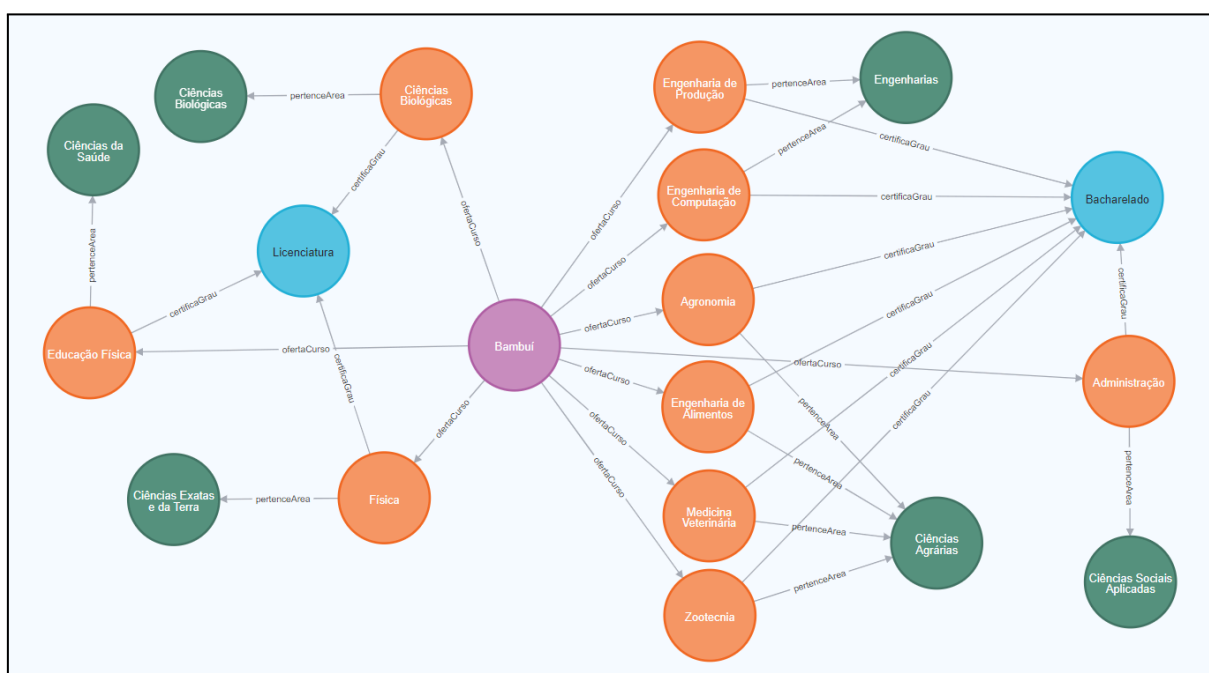
Figura 49: Grafo com as relações de um curso específico



Fonte: Elaborada pelo autor (2023)

A amplitude das possibilidades de melhoria na visualização dos dados com o formato de grafos também pode ser percebida na Figura 50, buscando evidenciar o rol de cursos ofertados por um Campus com base nos graus acadêmicos conferidos por eles e nas suas áreas de conhecimento. Neste exemplo, tem-se a estruturação do Campus Bambuí, onde são oferecidos dez cursos superiores, sendo que a maioria (quatro) está vinculada com a área de Ciências Agrárias. Quanto ao grau acadêmico, três cursos são de Licenciatura e sete de Bacharelado.

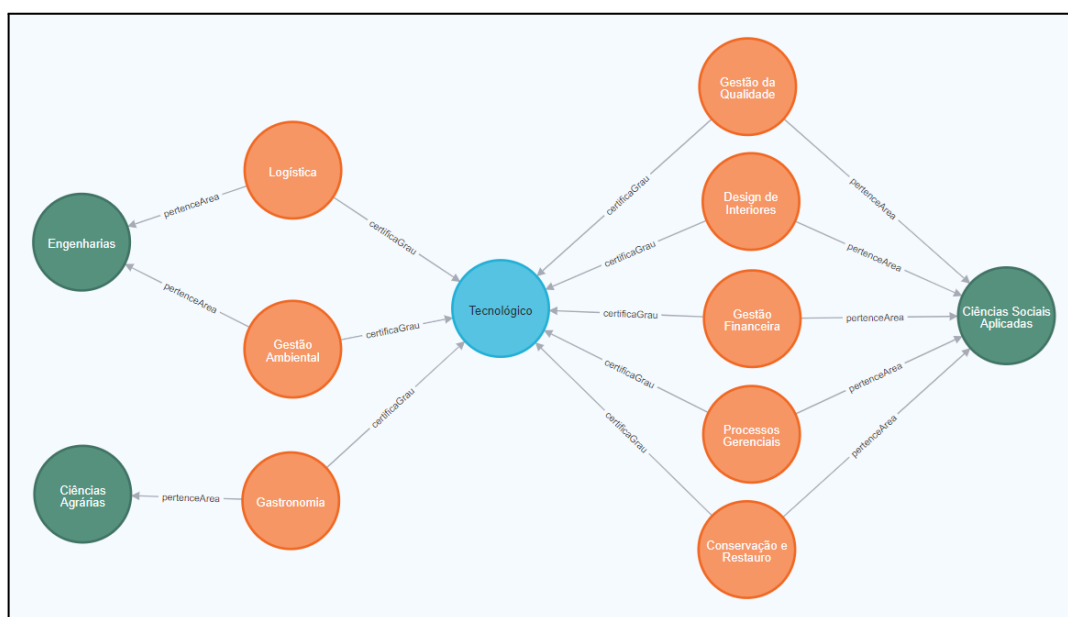
Figura 50: Grafo com os cursos ofertados por determinado Campus



Fonte: Elaborada pelo autor (2023)

O escopo do grafo foi, intencionalmente, expandido na Figura 51 para englobar os cursos de grau Tecnológico ofertados em todos os *campi* do IFMG, em conjunto com as suas áreas de conhecimento. Verifica-se considerável prevalência de cursos da área de Ciências Sociais Aplicadas, o que poderia ensejar uma discussão no âmbito acadêmico sobre a possibilidade de incrementar a diversidade de áreas contempladas nesta oferta em futuras autorizações para abertura de novos cursos.

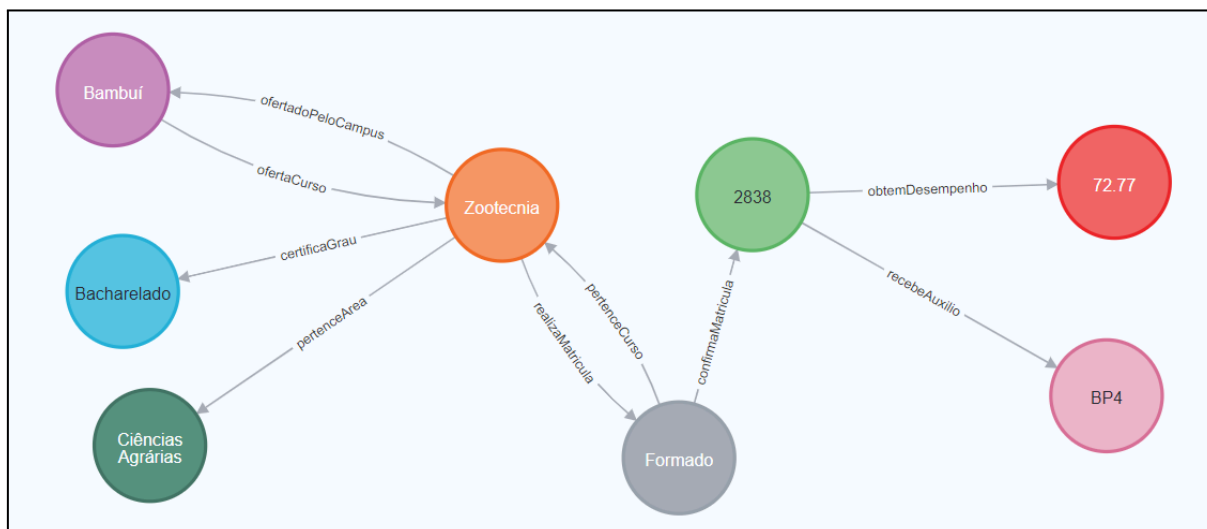
Figura 51: Grafo com os cursos de grau Tecnológico ofertados pelo IFMG



Fonte: Elaborada pelo autor (2023)

Na sequência, foram importados os registros que povoam as classes *Estudante*, *DesempenhoAcademico*, *Bolsista* (com a relação *recebeAuxilio*) e *Matricula* (com as relações *ingressoPor*, *pertenceCurso*, *confirmaMatricula* e *obtemDesempenho*). Ao final dessa atividade, o repositório do *Neo4j* contabilizava 41.088 nós e 94.509 relações. Não tendo o intuito de explorar a visualização no formato de grafos de todos os dados do repositório, mas tão somente para exemplificar como ela é útil para compreender acerca da estruturação dos nós e relações, além da instanciação dos registros, observa-se na Figura 52 toda a cadeia de relacionamentos de um estudante bolsista, cujo número de identificação é 2838. O seu coeficiente de rendimento geral foi de 72,77, ele recebeu auxílio socioeconômico no formato de Bolsa Permanência 4 (BP4), se formou no curso de Bacharelado em Zootecnia ofertado pelo Campus Bambuí e pertencente à área de Ciências Agrárias.

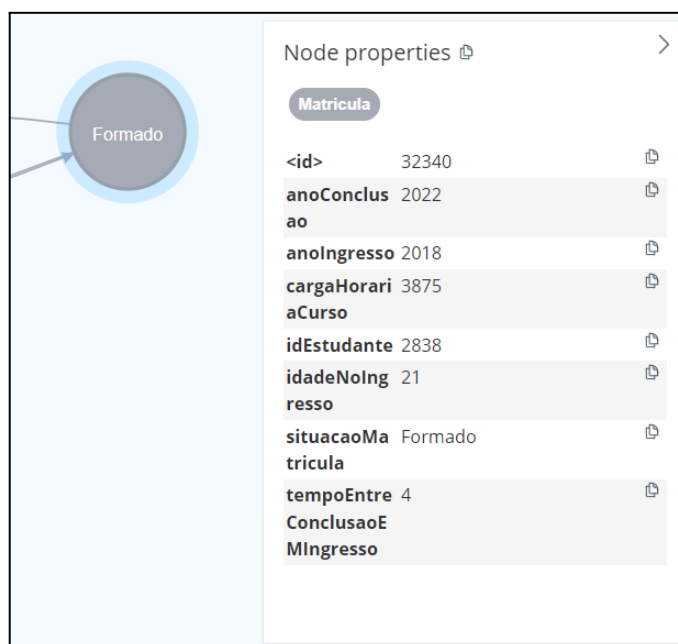
Figura 52: Grafo com as relações de um estudante bolsista



Fonte: Elaborada pelo autor (2023)

Embora essa visualização do grafo exiba apenas uma das propriedades dos dados de cada nó, todas aquelas constantes nos *scripts* de importação permanecem na estrutura do mesmo. No caso do software *Neo4j Browser*, é preciso selecionar um nó para que suas propriedades sejam exibidas em uma janela adicional. Na Figura 53 é possível notar a janela *Node properties* contendo as propriedades de dados do nó da classe *Matricula* de determinado estudante.

Figura 53: Propriedades de um nó selecionado



Fonte: Elaborada pelo autor (2023)

A ingestão dos dados no repositório em grafo do *Neo4j*, em alinhamento com a ontologia operacional construída anteriormente, promoveu o ambiente adequado para a execução de consultas que respondessem às questões de competência elencadas na Seção 4.3 e, por conseguinte, fornecessem os subsídios necessários para a realização da avaliação da ontologia.

4.3.3.4 Avaliação da ontologia

De acordo com o especificado no Capítulo 3 (Metodologia), a avaliação da qualidade da ontologia refere-se à verificação se a mesma atende aos requisitos propostos quando da sua concepção. Para isto, a tarefa inicial executada consistiu na elaboração de consultas escritas na linguagem *Cypher* para responder as questões de competência da ontologia por meio da recuperação das informações armazenadas no repositório em grafo do *Neo4j*.

A seguir são construídas as respostas para cada questão de competência, tendo como entrada a consulta escrita em *Cypher* e como saída um resultado quantitativo processado. Para aqueles resultados que envolvem comparações entre bolsistas e não bolsistas, foi verificado se a diferença numérica obtida era estatisticamente significativa com a aplicação do Teste T de Student. No repositório virtual do projeto em <https://github.com/educardosomelo/onto-educacional/tree/main> está disponível uma planilha eletrônica com o detalhamento de todos os testes conduzidos para validação estatística.

QC1: O desempenho acadêmico dos estudantes atendidos por auxílio socioeconômico é superior ao daqueles que não são atendidos?

Consulta escrita em *Cypher*:

```
MATCH (bol:Bolsista)-[:obtemDesempenho]->(des:DesempenhoAcademico)
WITH COUNT(bol) AS qtd, AVG(des.crGeral) AS cr_geral_medio, 'Bolsistas' AS
tipo
RETURN tipo, qtd AS qtd, ROUND(cr_geral_medio, 2) AS cr_geral_medio
UNION ALL
MATCH (est:Estudante)-[:obtemDesempenho]->(des:DesempenhoAcademico)
WITH COUNT(est) AS qtd, AVG(des.crGeral) AS cr_geral_medio, 'Não bolsistas'
AS tipo
RETURN tipo, qtd AS qtd, ROUND(cr_geral_medio, 2) AS cr_geral_medio;
```

Nesta tese, o indicador escolhido para medir o desempenho acadêmico é o coeficiente de rendimento (CR) do estudante obtido ao longo do curso. Conforme observa-se na Tabela 13, esta questão de competência envolveu a análise dos 10.263 estudantes que compõem o conjunto de dados da pesquisa, sendo que 2.142 (20,9%) são

bolsistas e 8.121 (79,1%) não receberam nenhum tipo de auxílio socioeconômico. A média do coeficiente de rendimento dos estudantes bolsistas é de 58,99, superior à média dos estudantes não bolsistas que é de 40,15. Trata-se de uma considerável diferença (18,84 pontos ou 32% superior) entre os dois grupos, indicando que, no contexto desta pesquisa, os estudantes que recebem auxílio socioeconômico possuem desempenho acadêmico superior em comparação com aqueles que não recebem.

Tabela 13: Resultados da consulta para a questão de competência 1

Grupo	Quantidade	CR geral médio
Bolsistas	2.142	58,99
Não bolsistas	8.121	40,15

Fonte: Elaborada pelo autor (2023)

Os resultados da validação estatística comprovam a significância da diferença do coeficiente de rendimento médio entre bolsistas e não bolsistas. É importante mencionar que o Teste T foi aplicado em quatro contextos diferentes, sendo que o primeiro comparou o CR dos bolsistas com o CR dos não bolsistas, mesmo com o evidente desbalanceamento entre esses grupos quando considerada a quantidade de estudantes em cada um. Visando reduzir a chance de influência deste desbalanceamento na validação, optou-se por criar três agrupamentos entre os registros dos estudantes não bolsistas, cada um possuindo a mesma quantidade do grupo de bolsistas (2.142); os CRs que compõem esses agrupamentos foram escolhidos aleatoriamente a partir do conjunto de registros de estudantes não bolsistas. A diferença entre os grupos comparados possui significância estatística ($p < 0,05$) em todos os contextos analisados, promovendo, assim, maior confiança no resultado encontrado para a questão de competência investigada (Tabela 14).

Tabela 14: Validação estatística dos resultados da questão de competência 1

Contexto	p	Significante
Bolsistas x Não bolsistas (desbalanceado)	< 0,001 (3,7841E-182)	Sim
Bolsistas x Não bolsistas 1	< 0,001 (1,7017E-110)	Sim
Bolsistas x Não bolsistas 2	< 0,001 (1,16E-114)	Sim
Bolsistas x Não bolsistas 3	< 0,001 (1,02623E-95)	Sim

Fonte: Elaborada pelo autor (2023)

QC2: Os estudantes que recebem auxílio socioeconômico evadem menos proporcionalmente do que aqueles que não recebem?

Consulta escrita em *Cypher*:

```
MATCH
(mat:Matricula)-[:confirmaMatricula]->(bol:Bolsista)-[:obtemDesempenho]->(des:DesempenhoAcademico)
WITH mat.situacaoMatricula AS situacao_matricula, COUNT(mat) AS qtd
WITH COLLECT({situacao_matricula: situacao_matricula, qtdSituacaoMatricula: qtd}) AS dados, SUM(qtd) AS total_situacao
UNWIND dados AS d
WITH d.situacao_matricula AS situacao_matricula_estudantes, 100.0 *
d.qtdSituacaoMatricula / total_situacao AS perc, d.qtdSituacaoMatricula AS
qtdSituacaoMatricula, 'Bolsistas' AS tipo, total_situacao AS total_situacao
WHERE situacao_matricula_estudantes = 'Evadido'
RETURN tipo, qtdSituacaoMatricula AS qtd_evadidos, total_situacao AS
total_do_grupo, ROUND(perc,1) AS proporcao
UNION ALL
MATCH
(mat:Matricula)-[:confirmaMatricula]->(est:Estudante)-[:obtemDesempenho]->(des:DesempenhoAcademico)
WITH mat.situacaoMatricula AS situacao_matricula, COUNT(mat) AS qtd
WITH COLLECT({situacao_matricula: situacao_matricula, qtdSituacaoMatricula: qtd}) AS dados, SUM(qtd) AS total_situacao
UNWIND dados AS d
WITH d.situacao_matricula AS situacao_matricula_estudantes, 100.0 *
d.qtdSituacaoMatricula / total_situacao AS perc, d.qtdSituacaoMatricula AS
qtdSituacaoMatricula, 'Não bolsistas' AS tipo, total_situacao AS
total_situacao
WHERE situacao_matricula_estudantes = 'Evadido'
RETURN tipo, qtdSituacaoMatricula AS qtd_evadidos, total_situacao AS
total_do_grupo, ROUND(perc,1) AS proporcao;
```

Analisando a Tabela 15, percebe-se que a proporção de estudantes bolsistas que evadiram (26,9%) é menor do que entre os não bolsistas (47,5%), permitindo inferir a contribuição positiva do recebimento do auxílio socioeconômico na taxa de evasão escolar. Quando investigada a diferença entre a proporção dos dois grupos, têm-se que a evasão dos bolsistas é 43,5% menor do que aquela dos não bolsistas. A aplicação do Teste T confirmou que existe significância estatística na diferença encontrada, com $p < 0,001$ (2,2E-16). Entretanto, mesmo este resultado entre os bolsistas enseja o planejamento de novas ações com vistas à redução da proporção de estudantes evadidos neste grupo, pois o dispêndio de dinheiro público é maior em função dos auxílios pagos.

Tabela 15: Resultados da consulta para a questão de competência 2

Grupo	Estudantes	Evadidos	Proporção
Bolsistas	2.142	577	26,9%
Não bolsistas	8.121	3.857	47,5%

Fonte: Elaborada pelo autor (2023)

QC3: A proporção de estudantes formados é maior entre aqueles que recebem auxílio socioeconômico?

Consulta escrita em *Cypher*:

```

MATCH
  (mat:Matricula)-[:confirmaMatricula]->(bol:Bolsista)-[:obtemDesempenho]->(des:DesempenhoAcademico)
WITH mat.situacaoMatricula AS situacao_matricula, COUNT(mat) AS qtd
WITH COLLECT({situacao_matricula: situacao_matricula, qtdSituacaoMatricula: qtd})
AS dados, SUM(qtd) AS total_situacao
UNWIND dados AS d
WITH d.situacao_matricula AS situacao_matricula_estudantes, 100.0 *
d.qtdSituacaoMatricula / total_situacao AS perc, d.qtdSituacaoMatricula AS
qtdSituacaoMatricula, 'Bolsistas' AS tipo, total_situacao AS total_situacao
WHERE situacao_matricula_estudantes = 'Formado'
RETURN tipo, qtdSituacaoMatricula AS qtd_formados, total_situacao AS
total_do_grupo, ROUND(perc,1) AS proporcao
UNION ALL
MATCH
  (mat:Matricula)-[:confirmaMatricula]->(est:Estudante)-[:obtemDesempenho]->(des:DesempenhoAcademico)
WITH mat.situacaoMatricula AS situacao_matricula, COUNT(mat) AS qtd
WITH COLLECT({situacao_matricula: situacao_matricula, qtdSituacaoMatricula:
qtd}) AS dados, SUM(qtd) AS total_situacao
UNWIND dados AS d
WITH d.situacao_matricula AS situacao_matricula_estudantes, 100.0 *
d.qtdSituacaoMatricula / total_situacao AS perc, d.qtdSituacaoMatricula AS
qtdSituacaoMatricula, 'Não bolsistas' AS tipo, total_situacao AS
total_situacao
WHERE situacao_matricula_estudantes = 'Formado'
RETURN tipo, qtdSituacaoMatricula AS qtd_formados, total_situacao AS
total_do_grupo, ROUND(perc,1) AS proporcao;

```

O resultado da terceira questão de competência fornece indícios similares aos obtidos nas questões anteriores sobre o papel do auxílio socioeconômico no apoio aos

estudantes do IFMG. Verifica-se na Tabela 16 que a proporção de estudantes bolsistas que se formaram (11,2%) é duas vezes superior à proporção dos não bolsistas (5,3%). Ressalta-se, para efeitos de caracterização do contexto, que os estudantes formados correspondem a apenas 6,5% das instâncias armazenadas no repositório, as quais englobam matrículas iniciadas em 2018 e com movimentação até o segundo semestre letivo de 2022, indicando assim possível dificuldade dos estudantes do IFMG finalizarem seus cursos no prazo mínimo regulamentar de três e meio, quatro ou cinco anos, seja por evasão ou retenção.

Tabela 16: Resultados da consulta para a questão de competência 3

Grupo	Estudantes	Formados	Proporção
Bolsistas	2.142	239	11,2%
Não bolsistas	8.121	431	5,3%

Fonte: Elaborada pelo autor (2023)

QC4: Existe relação entre a evasão, desempenho acadêmico e as características pessoais dos estudantes bolsistas e não bolsistas, tais como gênero e cor?

Consulta escrita em *Cypher* investigando as relações com o **gênero**:

```
CALL {
MATCH (mat:Matricula {situacaoMatricula:
'Evadido'})-[:confirmaMatricula]->(est:Estudante)
MATCH (est:Estudante)-[:obtemDesempenho]->(des:DesempenhoAcademico)
WITH est.genero AS genero, COUNT(mat) AS qtd, SUM(des.crGeral) AS cr_geral
WITH COLLECT({genero: genero, qtdGenero: qtd, cr_geral: cr_geral}) AS dados,
SUM(qtd) AS total_evadidos
UNWIND dados AS d
WITH d.genero AS genero, 100.0 * d.qtdGenero / total_evadidos AS perc,
d.qtdGenero AS qtdGenero, d.cr_geral / d.qtdGenero AS cr_geral_medio
RETURN genero, 'Não bolsistas' AS tipo, qtdGenero AS qtd, ROUND(perc,1) AS
percentual_entre_evadidos, ROUND(cr_geral_medio, 2) AS cr_geral_medio
UNION ALL
MATCH (mat:Matricula {situacaoMatricula:
'Evadido'})-[:confirmaMatricula]->(bol:Bolsista)
MATCH (bol:Bolsista)-[:obtemDesempenho]->(des:DesempenhoAcademico)
WITH bol.genero AS genero, COUNT(mat) AS qtd, SUM(des.crGeral) AS cr_geral
WITH COLLECT({genero: genero, qtdGenero: qtd, cr_geral: cr_geral}) AS dados,
SUM(qtd) AS total_evadidos
```

```

UNWIND dados AS d WITH d.genero AS genero, 100.0 * d.qtdGenero /
total_evadidos AS perc, d.qtdGenero AS qtdGenero, d.cr_geral / d.qtdGenero AS
cr_geral_medio
RETURN genero, 'Bolsistas' AS tipo, qtdGenero AS qtd, ROUND(perc,1) AS
percentual_entre_evadidos, ROUND(cr_geral_medio, 2) AS cr_geral_medio
}
RETURN genero, tipo, qtd, percentual_entre_evadidos, cr_geral_medio
ORDER BY genero, tipo;

```

A quarta questão de competência envolveu a análise da relação entre três características pessoais dos estudantes (gênero, cor e se a naturalidade é na mesma cidade onde o Campus de vínculo está localizado) com os indicadores de desempenho acadêmico e evasão. Os resultados da propriedade *gênero* são expostos na Tabela 17, evidenciando novamente diferenças entre os bolsistas e não bolsistas nos dois indicadores analisados. Enquanto no grupo de bolsistas a evasão foi maior entre o gênero feminino (54,1% a 45,9%), situação inversa ocorreu no grupo de não bolsistas, tendo o gênero masculino maior percentual de evadidos (57,6% a 42,4%). Estudantes bolsistas do gênero feminino apresentaram coeficiente de rendimento acadêmico médio superior a quase o dobro do obtido pelas não bolsistas (42,41 a 23,20). No caso dos estudantes do gênero masculino, o CR dos bolsistas (38,36) também é amplamente maior do que o dos não bolsistas (21,52).

Tabela 17: Resultados da consulta para a questão de competência 4 (gênero)

Gênero	Grupo	Evadidos	% entre evadidos	CR geral médio
Feminino	Bolsistas	312	54,1%	42,41
	Não bolsistas	1.637	42,4%	23,20
Masculino	Bolsistas	265	45,9%	38,36
	Não bolsistas	2.220	57,6%	21,52

Fonte: Elaborada pelo autor (2023)

Para validar se essas diferenças entre os CRs eram estatisticamente significativas, o Teste T foi aplicado em oito contextos, sendo quatro para cada gênero: o primeiro entendendo como indiferente o desbalanceamento na quantidade de estudantes evadidos em cada grupo e os três seguintes igualando as quantidades de registros nos grupos (em função da quantidade de bolsistas) para reduzir os efeitos deste desbalanceamento. Os CRs dos agrupamentos de não bolsistas nomeados como 1, 2 e 3 foram escolhidos aleatoriamente. A Tabela 18 disponibiliza os resultados que indicam significância estatística ($p < 0,05$) em todos os contextos verificados.

Tabela 18: Validação estatística dos resultados da questão de competência 4 (gênero)

Gênero	Contexto	p	Significante
Feminino	Bolsistas x Não bolsistas (desb.)	< 0,001 (1,69385E-29)	Sim
	Bolsistas x Não bolsistas 1	< 0,001 (1,96996E-20)	Sim
	Bolsistas x Não bolsistas 2	< 0,001 (4,25224E-23)	Sim
	Bolsistas x Não bolsistas 3	< 0,001 (1,20038E-36)	Sim
Masculino	Bolsistas x Não bolsistas (desb.)	< 0,001 (1,42979E-21)	Sim
	Bolsistas x Não bolsistas 1	< 0,001 (2,83246E-12)	Sim
	Bolsistas x Não bolsistas 2	< 0,001 (5,93722E-17)	Sim
	Bolsistas x Não bolsistas 3	< 0,001 (1,85208E-24)	Sim

Fonte: Elaborada pelo autor (2023)

Consulta escrita em *Cypher* investigando as relações com a **cor**: para a recuperação dos dados relacionados com a cor dos estudantes, foi necessário apenas substituir no comando *Cypher* criado anteriormente a propriedade *gênero* pela propriedade *cor*.

Verifica-se na Tabela 19 que quase a metade (48,9%) dos estudantes evadidos bolsistas se declarou como sendo da cor parda, 34,8% da cor branca, 14,9% da cor preta e o restante das cores amarelo e indígena. Os percentuais são muito próximos entre os estudantes evadidos não bolsistas. O coeficiente de rendimento acadêmico médio dos bolsistas foi superior ao dos não bolsistas em todas as classificações de cor.

Tabela 19: Resultados da consulta para a questão de competência 4 (cor)

Cor	Grupo	Evadidos	% entre evadidos	CR geral médio
Amarela	Bolsistas	6	1,0%	21,27
	Não bolsistas	41	1,1%	17,99
Branca	Bolsistas	201	34,8%	43,37
	Não bolsistas	1.420	36,8%	23,90
Indígena	Bolsistas	2	0,3%	30,09
	Não bolsistas	4	0,1%	23,76
Parda	Bolsistas	282	48,9%	39,80
	Não bolsistas	1.756	45,5%	21,68
Preta	Bolsistas	86	14,9%	38,23
	Não bolsistas	636	16,5%	20,08

Fonte: Elaborada pelo autor (2023)

Em virtude da maior representatividade das cores branca, parda e preta entre os estudantes evadidos, elas foram selecionadas para participação na validação estatística das diferenças entre o CR dos bolsistas e não bolsistas. Seguindo a mesma padronização do Teste T aplicado na QC1 e na primeira parte dessa QC4, houve a definição de quatro contextos para cada uma dessas três cores, totalizando 12 contextos validados. O primeiro contexto de uma cor é caracterizado pelo desbalanceamento na quantidade de registros entre bolsistas e não bolsistas, enquanto nos outros três a quantidade de registros de não bolsistas sempre foi igual a de bolsistas. A execução dos testes, cujos resultados estão dispostos na Tabela 20, mostrou que existe significância estatística na diferença dos coeficientes de rendimento médio em todos os doze contextos validados.

Tabela 20: Validação estatística da questão de competência 4 (cor)

Cor	Contexto	<i>p</i>	Significante
Branca	Bolsistas x Não bolsistas (desb.)	< 0,001 (2,92611E-21)	Sim
	Bolsistas x Não bolsistas 1	< 0,001 (4,35677E-09)	Sim
	Bolsistas x Não bolsistas 2	< 0,001 (5,27966E-13)	Sim
	Bolsistas x Não bolsistas 3	< 0,001 (4,32216E-17)	Sim
Parda	Bolsistas x Não bolsistas (desb.)	< 0,001 (1,02442E-25)	Sim
	Bolsistas x Não bolsistas 1	< 0,001 (5,36791E-21)	Sim
	Bolsistas x Não bolsistas 2	< 0,001 (8,99695E-09)	Sim
	Bolsistas x Não bolsistas 3	< 0,001 (2,62705E-24)	Sim
Preta	Bolsistas x Não bolsistas (desb.)	< 0,001 (5,55273E-08)	Sim
	Bolsistas x Não bolsistas 1	< 0,001 (6,40206E-05)	Sim
	Bolsistas x Não bolsistas 2	< 0,001 (1,05464E-09)	Sim
	Bolsistas x Não bolsistas 3	< 0,001 (2,90184E-06)	Sim

Fonte: Elaborada pelo autor (2023)

QC5: Existem variações no desempenho acadêmico e na taxa de evasão de acordo com o tipo de Bolsa Permanência destinada aos estudantes?

Consulta escrita em *Cypher*:

```
MATCH (bol:Bolsista)-[:recebeAuxilio]->(aux:AuxilioSocioeconomico)
MATCH (bol:Bolsista)-[:obtemDesempenho]->(des:DesempenhoAcademico)
WITH aux, COUNT(bol) AS qtd, AVG(des.crGeral) AS cr_geral_medio
```

```

WITH COLLECT({tipoBolsa: aux.nome, qtdBolsistas: qtd, crGeralMedio:
cr_geral_medio}) AS dados, SUM(qtd) AS total_bolsistas
UNWIND dados AS d
MATCH (mat:Matricula {situacaoMatricula:
'Evadido'})-[:confirmaMatricula]->(bol:Bolsista)-[:recebeAuxilio]->(aux:Auxil
ioSocioeconomico {nome: d.tipoBolsa})
WITH aux, COUNT(mat) AS qtd_evadidos, d.tipoBolsa AS tipoBolsa,
d.qtdBolsistas AS qtdBolsistas, d.crGeralMedio AS crGeralMedio,
total_bolsistas AS totalBolsistas
WITH COLLECT({tipoBolsa: tipoBolsa, qtdBolsistas: qtdBolsistas, crGeralMedio:
crGeralMedio, totalBolsistas: totalBolsistas, qtdEvadidos: qtd_evadidos}) AS
dados2
UNWIND dados2 AS e
WITH e.tipoBolsa AS tipo_bolsa, e.qtdBolsistas AS qtd_bolsistas,
e.crGeralMedio AS cr_geral_medio, e.qtdEvadidos AS qtd_evadidos, 100.0 *
e.qtdEvadidos / e.qtdBolsistas AS taxa_evasao_tipo_bolsa
RETURN tipo_bolsa, qtd_bolsistas AS qtd_bolsistas, ROUND(cr_geral_medio,2) AS
cr_geral_medio, qtd_evadidos AS qtd_evadidos, ROUND(taxa_evasao_tipo_bolsa,1)
AS taxa_evasao_tipo_bolsa
ORDER BY tipo_bolsa;

```

A distribuição da Bolsa Permanência (BP) no IFMG é materializada com o pagamento mensal de um auxílio que pode ser de quatro tipos: BP1 (R\$400,00), BP2 (R\$300,00), BP3 (R\$200,00) e BP4 (R\$150,00). São pagas dez parcelas anuais para os estudantes que cumprem os requisitos do Programa de Assistência Estudantil do IFMG. No caso dos estudantes evadidos, em média são feitos três pagamentos até que ocorra a efetivação da saída da instituição. A quinta questão de competência objetivou fornecer mais detalhes sobre os estudantes que, mesmo recebendo auxílio socioeconômico, evadiram de seus cursos.

De acordo com a Tabela 21, BP4 é o tipo de bolsa mais distribuído aos estudantes, seguido por BP3, BP2 e, por fim, BP1, em uma ordem inversamente proporcional ao valor pago por cada uma delas. Essa sequência, contudo, não se repete ao ser analisada a taxa de evasão por tipo de bolsa; enquanto a menor taxa foi observada entre os estudantes que recebem a BP4, taxas próximas e superiores caracterizam a evasão de quem é beneficiado com os outros três tipos de bolsa. Os dados indicam que 60% dos bolsistas que evadiram não conseguiram se manter no curso, mesmo recebendo valores maiores de auxílio socioeconômico. Quando o coeficiente de rendimento acadêmico é apreciado, observa-se uma situação na qual os valores das bolsas diminuem (BP1 para BP4) enquanto as médias aumentam (46,58 para 61,67). Uma hipótese para que a menor média advenha justamente dos estudantes beneficiados com o tipo de bolsa de maior valor (BP1) é que trata-se de

indivíduos com alta vulnerabilidade socioeconômica que convivem em contextos familiares marcados por situações que dificultam rotinas de estudos.

Tabela 21: Resultados da consulta para a questão de competência 5

Tipo de bolsa	Bolsistas	Evadidos	Taxa de evasão	CR geral médio
BP1	61	18	29,5%	46,58
BP2	347	107	30,8%	55,57
BP3	693	218	31,5%	57,77
BP4	1.041	234	22,5%	61,67

Fonte: Elaborada pelo autor (2023)

Ao considerarmos o cenário do IFMG exposto na questão de competência e o fato de que os bolsistas receberam, em média, três parcelas do auxílio antes de evadir, temos que a instituição distribuiu, entre 2018 e 2022, cerca de R\$ 354.000,00 (trezentos e cinquenta e quatro mil reais) para estudantes que não traduziram diretamente o investimento em um maior quantitativo de formandos.

QC6: Como é a distribuição da evasão considerando as áreas de conhecimentos dos cursos, além da taxa de evasão de cada área?

Consulta escrita em *Cypher*:

```
MATCH
(mat:Matricula)-[:pertenceCurso]->(cur:Curso)-[:pertenceArea]->(are:AreaDeConhecimento)
WITH are, COUNT(mat) AS qtd
WITH COLLECT({area:are.nome, qtdEstudantes: qtd}) AS dados
UNWIND dados AS d
MATCH (mat2:Matricula {situacaoMatricula: 'Evadido'})-[:pertenceCurso]->(cur2:Curso)-[:pertenceArea]->(are2:AreaDeConhecimento {nome: d.area})
WITH are2, COUNT(mat2) AS qtd_evadidos, d.qtdEstudantes AS qtd_estudantes
WITH COLLECT({area:are2.nome, qtdEvadidos: qtd_evadidos, qtdEstudantes: qtd_estudantes}) AS dados2, SUM(qtd_evadidos) AS total_evadidos
UNWIND dados2 AS e
WITH e.area AS area_conhecimento, 100.0 * e.qtdEvadidos / total_evadidos AS perc, e.qtdEstudantes AS total_estudantes, e.qtdEvadidos AS qtd_evadidos, 100.0 * e.qtdEvadidos / e.qtdEstudantes AS taxa_evasao_area
```

```

RETURN area_conhecimento, qtd_evadidos AS qtd_evadidos, ROUND(perc,1) AS
percentual_entre_evadidos, total_estudantes AS total_estudantes_area,
ROUND(taxa_evasao_area,1) AS taxa_evasao_area
ORDER BY area_conhecimento;

```

Quando a evasão é visualizada sob a ótica das áreas de conhecimentos dos cursos (Tabela 22), o primeiro ponto que sinaliza atenção é o fato de que a taxa de evasão em todas as áreas é superior a 38%, com exceção das Ciências da Saúde, área com o menor quantitativo de estudantes na instituição. Os cursos da área de Ciências Exatas e da Terra formam a mais alta proporção de evasão do IFMG (50,6%), embora não tenham sequer a metade do número de estudantes de cursos da área de Engenharias e ciente de que os cursos de ambas as áreas compartilham, de certa maneira, uma formação introdutória mais complexa. Também merece exposição o resultado obtido pela área de Ciências Sociais Aplicadas, pois mesmo tendo o segundo maior quantitativo de estudantes, obteve taxa de evasão de 38%. Os dados demonstram que é preciso conduzir, por parte da administração acadêmica, estudos mais aprofundados em cursos de determinadas áreas de conhecimento para entender os fatores internos e externos que subsidiam a concretização da evasão pelos estudantes.

Tabela 22: Resultados da consulta para a questão de competência 6

Área de conhecimento	Estudantes da área	Evadidos	Taxa de evasão
Ciências Agrárias	1.214	465	38,3%
Ciências Biológicas	335	158	47,2%
Ciências Exatas e da Terra	1.524	771	50,6%
Ciências Humanas	425	178	41,9%
Ciências Sociais Aplicadas	2.723	1.036	38,0%
Ciências da Saúde	41	11	26,8%
Engenharias	3.869	1.764	45,6%
Linguística Letras e Artes	132	51	38,6%

Fonte: Elaborada pelo autor (2023)

QC7: Como é a distribuição da evasão considerando o grau certificado pelos cursos, além da taxa de evasão de cada grau?

Consulta escrita em *Cypher*:

```

MATCH
(mat:Matricula)-[:pertenceCurso]->(cur:Curso)-[:certificaGrau]->(grau:GrauAcademico)
WITH grau, COUNT(mat) AS qtd
WITH COLLECT({grau:grau.nome, qtdEstudantes: qtd}) AS dados
UNWIND dados AS d
MATCH (mat2:Matricula {situacaoMatricula:
'Evadido'})-[:pertenceCurso]->(cur2:Curso)-[:certificaGrau]->(grau2:GrauAcademico {nome: d.grau})
WITH grau2, COUNT(mat2) AS qtd_evadidos, d.qtdEstudantes AS qtd_estudantes
WITH COLLECT({grau:grau2.nome, qtdEvadidos: qtd_evadidos, qtdEstudantes:
qtd_estudantes}) AS dados2, SUM(qtd_evadidos) AS total_evadidos
UNWIND dados2 AS e
WITH e.grau AS grau_academico, 100.0 * e.qtdEvadidos / total_evadidos AS
perc, e.qtdEstudantes AS total_estudantes, e.qtdEvadidos AS qtd_evadidos,
100.0 * e.qtdEvadidos / e.qtdEstudantes AS taxa_evasao_grau
RETURN grau_academico, qtd_evadidos AS qtd_evadidos, ROUND(perc,1) AS
percentual_entre_evadidos, total_estudantes AS total_estudantes_grau,
ROUND(taxa_evasao_grau,1) AS taxa_evasao_grau
ORDER BY grau_academico;

```

A sétima questão de competência permite compreender melhor a evasão a partir do agrupamento dos cursos conforme seu grau acadêmico. Verifica-se, na Tabela 23, considerável discrepância entre a quantidade de estudantes em cada grau, sendo que a soma das matrículas dos cursos de Licenciatura e Tecnológico não corresponde sequer à metade do número dos estudantes de Bacharelado. Não obstante apresentar a maior quantidade de estudantes evadidos, esse grau acadêmico obteve a menor taxa de evasão entre os três grupos analisados. Sugere-se a realização, pelo IFMG, de estudos específicos para coletar novos dados que permitam melhorar o entendimento sobre os motivos que levam mais da metade dos estudantes dos cursos de Licenciatura a encerrarem seus vínculos e, assim, incrementarem o resultado da evasão deste grau acadêmico.

Tabela 23: Resultados da consulta para a questão de competência 7

Grau acadêmico	Estudantes do grau	Evadidos	Taxa de evasão
Bacharelado	7.106	2.897	40,8%
Licenciatura	1.601	813	50,8%
Tecnológico	1.556	724	46,5%

Fonte: Elaborada pelo autor (2023)

QC8: Como é a distribuição da evasão considerando a forma de ingresso do estudante ao curso, além da taxa de evasão de cada forma?

Consulta escrita em *Cypher*:

```
MATCH (mat:Matricula)-[:ingressoPor]->(fin:FormaDeIngresso)
WITH fin, COUNT(mat) AS qtd
WITH COLLECT({forma: fin.nome, qtdEstudantes: qtd}) AS dados
UNWIND dados AS d
MATCH (mat2:Matricula {situacaoMatricula:
'Evadido'})-[:ingressoPor]->(fin2:FormaDeIngresso {nome: d.forma})
WITH fin2, COUNT(mat2) AS qtd_evadidos, d.qtdEstudantes AS qtd_estudantes
WITH COLLECT({forma: fin2.nome, qtdEvadidos: qtd_evadidos, qtdEstudantes:
qtd_estudantes}) AS dados2, SUM(qtd_evadidos) AS total_evadidos
UNWIND dados2 AS e
WITH e.forma AS forma_ingresso, 100.0 * e.qtdEvadidos / total_evadidos AS
perc, e.qtdEstudantes AS total_estudantes, e.qtdEvadidos AS qtd_evadidos,
100.0 * e.qtdEvadidos / e.qtdEstudantes AS taxa_evasao_forma
RETURN forma_ingresso, qtd_evadidos AS qtd_evadidos, ROUND(perc,1) AS
percentual_entre_evadidos, total_estudantes AS total_estudantes_forma,
ROUND(taxa_evasao_forma,1) AS taxa_evasao_forma
ORDER BY forma_ingresso;
```

A oitava e última questão de competência aborda a evasão sob a ótica da forma de ingresso ao curso pelo estudante. Verificando a Tabela 24, é possível perceber a considerável diferença no quantitativo de estudantes que ingressaram via SISU em comparação com as outras três formas. Os ingressos por Transferência externa e Obtenção de novo título são pouco expressivos em termos de quantidade, entretanto chama a atenção o fato da taxa de evasão entre os estudantes que buscam um novo título ser quase o dobro (53,1% a 28,6%) da calculada para aqueles que vieram transferidos de outras instituições. Seria interessante uma investigação mais detalhada, pela gestão acadêmica do IFMG, acerca de quais fatores têm contribuído para a elevada evasão entre os ingressantes via SISU, pois trata-se do principal meio de entrada dos estudantes na instituição. Eventualmente, ações como acolhimento mais próximo de estudantes oriundos de cidades distantes, nivelamento em conteúdos específicos de áreas básicas e adequada divulgação/caracterização dos cursos ofertados, dentre outras, podem se mostrar positivas e dificultar a ocorrência da evasão.

Tabela 24: Resultados da consulta para a questão de competência 8

Forma de ingresso	Estudantes da forma	Evadidos	Taxa de evasão
Obtenção de novo título	392	208	53,1%
SISU	7.020	3.139	44,7%
Transferência externa	388	111	28,6%
Vestibular	2.463	976	39,6%

Fonte: Elaborada pelo autor (2023)

Assim como mencionado anteriormente, a utilização de questões de competência tornou mais clara a abrangência da ontologia elaborada, indicando o escopo de sua atuação e, inclusive, definindo indiretamente aspectos que não fazem parte do domínio construído. O desenvolvimento do ambiente tecnológico que suportou a ingestão dos dados organizados pela ontologia foi de grande valia para que a transformação das questões de competência em consultas escritas com uma sintaxe específica produzisse resultados tangíveis e interpretáveis e, adicionalmente, colaborasse para avaliar a adequação da ontologia em relação a seus requisitos.

Como complemento à avaliação da ontologia, a criação de uma matriz de rastreabilidade (Quadro 6) vinculou a cada questão de competência aqueles elementos da ontologia utilizados na sua resposta, sejam conceitos, relações ou propriedades de dados. Dentre todas as classes definidas na ontologia, apenas a nomeada como *Campus* não foi envolvida por alguma questão de competência, especialmente devido ao fato de que a grande quantidade de unidades no IFMG tornaria a resposta de grande extensão. A análise sobre a evasão escolar, desempenho acadêmico e recebimento de auxílio socioeconômico por Campus é descrita na Seção 4.4, pois o *dashboard* gerencial construído contempla uma visualização dos dados agrupados por essa propriedade. A classe *Pessoa* não foi instanciada diretamente no repositório, apenas por meio de sua especialização via classe *Estudante*, por isso também não foi listada na matriz de rastreabilidade. Houve baixo emprego das propriedades dos dados das classes nas questões de competência, com destaque para aquelas pertencentes à classe *DesempenhoAcademico*; neste caso em particular, a motivação se remete à definição do coeficiente de rendimento geral como métrica para medir o desempenho acadêmico dos estudantes. No *dashboard* gerencial buscou-se utilizar mais propriedades de dados para integrar outros elementos da ontologia e servir de suporte para a formação de novos conhecimentos.

Quadro 6: Matriz de rastreabilidade das questões de competência da ontologia

Questão de competência	Conceitos, relações ou propriedades
1) O desempenho acadêmico dos estudantes atendidos por auxílio socioeconômico é superior ao daqueles que não são atendidos?	Bolsista, obtemDesempenho, DesempenhoAcademico, crGeral, Estudante
2) Os estudantes que recebem auxílio socioeconômico evadem menos proporcionalmente do que aqueles que não recebem?	Matricula, confirmaMatricula, Bolsista, obtemDesempenho, DesempenhoAcademico, situacaoMatricula, Estudante
3) A proporção de estudantes formados é maior entre aqueles que recebem auxílio socioeconômico?	Matricula, confirmaMatricula, Bolsista, obtemDesempenho, DesempenhoAcademico, situacaoMatricula, Estudante
4) Existe relação entre a evasão, desempenho acadêmico e as características pessoais dos estudantes bolsistas e não bolsistas, tais como gênero e cor?	Matricula, situacaoMatricula, confirmaMatricula, Estudante, obtemDesempenho, DesempenhoAcademico, genero, cor, crGeral, Bolsista
5) Existem variações no desempenho acadêmico e na taxa de evasão de acordo com o tipo de Bolsa Permanência destinada aos estudantes?	Bolsista, recebeAuxilio, AuxilioSocioeconomico, obtemDesempenho, DesempenhoAcademico, crGeral, nome, situacaoMatricula
6) Como é a distribuição da evasão considerando as áreas de conhecimentos dos cursos, além da taxa de evasão de cada área?	Matricula, pertenceCurso, Curso, pertenceArea, AreaDeConhecimento, nome, situacaoMatricula
7) Como é a distribuição da evasão considerando o grau certificado pelos cursos, além da taxa de evasão de cada grau?	Matricula, pertenceCurso, Curso, certificaGrau, GrauAcademico, nome, situacaoMatricula
8) Como é a distribuição da evasão considerando a forma de ingresso do estudante ao curso, além da taxa de evasão de cada forma?	Matricula, ingressoPor, FormaDeIngresso, nome, situacaoMatricula

Fonte: Elaborado pelo autor (2023)

O artefato produzido sob a forma de uma matriz de rastreabilidade mostra como as questões de competência estão alinhadas com a ontologia, permitindo afirmar que elas cobrem, adequadamente, quase a totalidade do domínio envolvido. Constata-se, ainda, a inexistência de excessos conceituais no âmbito da Onto-Educacional, pois todos os conceitos que fazem parte do seu escopo contribuem para a organização dos dados recebidos de diferentes fontes.

Nas próximas seções são apresentados os resultados das análises realizadas com o intuito de cumprir os objetivos gerais da tese.

4.4 Análise das relações entre os construtos

Para que as relações entre os construtos fosse analisada sob um ponto de vista complementar ao proporcionado pelos resultados das questões de competência, foi prototipado um *dashboard* gerencial com recursos de visualização de informações abrangendo todos os dados integrados com o suporte conceitual da Onto-Educacional. O

dashboard contém nove abas, assim organizadas: 1) Apresentação do artefato; 2) Indicadores consolidados; 3) Relação entre os três construtos por campus; 4) Evasão por curso/turma; 5) Bolsistas, evadidos e formados por gênero e cor; 6) Evasão e desempenho acadêmico por tipo de bolsa; 7) Desempenho acadêmico por campus/curso/turma/situação de matrícula; 8) Desempenho acadêmico consolidado por campus (parte 1); 9) Desempenho acadêmico consolidado por campus (parte 2). A versão final do *dashboard* está disponível no link:

<https://app.powerbi.com/view?r=eyJrjoiMzg2NjM3NjUtODg1NC00ZGY5LTk0MDMtNTA5MmZiZTFiZDM0liwidCI6IjNhNjI3ZDMwLWZhN2YtNDI3NS04OGJILWQzYTk0MTFIODExNCJ9&pageName=ReportSection>.

A primeira aba, reproduzida na Figura 54, serve para apresentar as informações gerais do *dashboard*, bem como caracterizá-lo no âmbito da pesquisa conduzida na tese.

Figura 54: Primeira aba do *dashboard*


Onto-Educacional

A Onto-Educacional é uma ontologia elaborada para recuperar informações sobre a distribuição de auxílios socioeconômicos no formato de bolsa permanência em programas da assistência estudantil do IFMG para analisar sua contribuição na redução da evasão escolar e no aumento do desempenho acadêmico dos estudantes de cursos superiores presenciais.


O escopo das análises envolve dados de estudantes de cursos presenciais de graduação ofertados pelos *campi* do IFMG entre o primeiro semestre letivo de 2018 e o segundo semestre letivo de 2022.

A proposta deste *dashboard* é consolidar as informações coletadas e apresentá-las com o apoio de gráficos e tabelas para que os usuários consigam obter novos conhecimentos envolvendo os construtos envolvidos na pesquisa.

A ontologia e este artefato no formato de *dashboard* foram construídos no contexto da tese de Doutorado do aluno Eduardo Cardoso Melo, orientado pelo professor Dr. Marcello Peixoto Bax, do Programa de Pós-Graduação em Gestão & Organização do Conhecimento (PPGGOC) da Universidade Federal de Minas Gerais (UFMG).



Programa de Pós Graduação
GESTÃO & ORGANIZAÇÃO DO
CONHECIMENTO



INSTITUTO FEDERAL
Minas Gerais

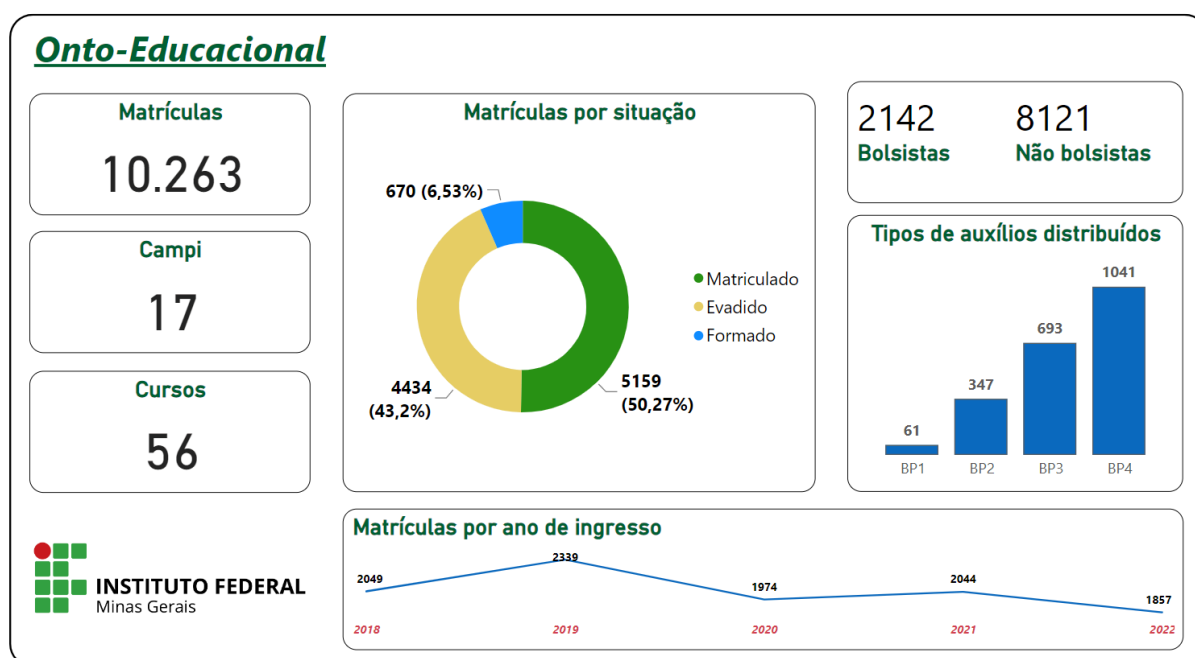
Fonte: Elaborada pelo autor (2024)

A segunda aba exibe indicadores consolidados de acordo com os dados armazenados no repositório em grafos do *Neo4j*. Observa-se, na Figura 55, que o conjunto total de dados contempla 10.263 registros de matrículas em 56 cursos de estudantes de 17 diferentes *campi*. A quantidade de registros varia em função da situação de cada matrícula, sendo 5.159 (50,3%) matriculados, 4.434 (43,2%) evadidos e somente 670 (6,5%) formados. Quanto à atribuição de bolsas, 2.142 (20,9%) são matrículas de bolsistas e 8.121 (79,1%) de não bolsistas. A distribuição de matrículas por tipo de auxílio segue uma relação

inversamente proporcional entre o valor de cada bolsa e a quantidade de bolsistas, isto é, quanto maior o valor da bolsa, menor é o número de estudantes que a recebem.

Considerando o quantitativo de estudantes e de bolsistas no âmbito da instituição, esse resultado pode indicar que, dada a insuficiência de recursos financeiros para atender toda a demanda, os gestores optaram por contemplar um número maior de estudantes com um valor menor de auxílio, a despeito da sua insuficiência para aplacar os efeitos da vulnerabilidade. A aba exibe, ainda, a distribuição das matrículas por ano de ingresso na instituição, onde observa-se forte diminuição em 2022 (cerca de 21% em comparação com 2019), possivelmente como um reflexo da pandemia de COVID-19.

Figura 55: Segunda aba do *dashboard*



Fonte: Elaborada pelo autor (2024)

O foco da terceira aba é servir de suporte para análise da relação entre os construtos em cada campus do IFMG. Por se tratar de uma listagem composta de muitos registros, os dados são apresentados por meio da Tabela 25, ao invés de uma figura com a aba. Nota-se que, com exceção do Campus Ibirité, a taxa de evasão entre os bolsistas é menor do que a obtida entre os não bolsistas, sendo que nos *campi* Arcos, Ouro Preto, Piumhi e Ponte Nova, essa diferença é superior a 30%. Esse resultado está de acordo com o obtido na segunda questão de competência da ontologia. O coeficiente de rendimento médio dos bolsistas é superior ao dos não bolsistas em todos os *campi*, exceto o de Ibirité, estando de acordo com o encontrado na resposta da primeira questão de competência. Tais números, mesmo que detalhados por campus, indicam novamente o impacto do recebimento de

auxílio socioeconômico na evasão escolar e no desempenho acadêmico dos estudantes de cursos de graduação do IFMG.

Tabela 25: Relação entre os três construtos por campus

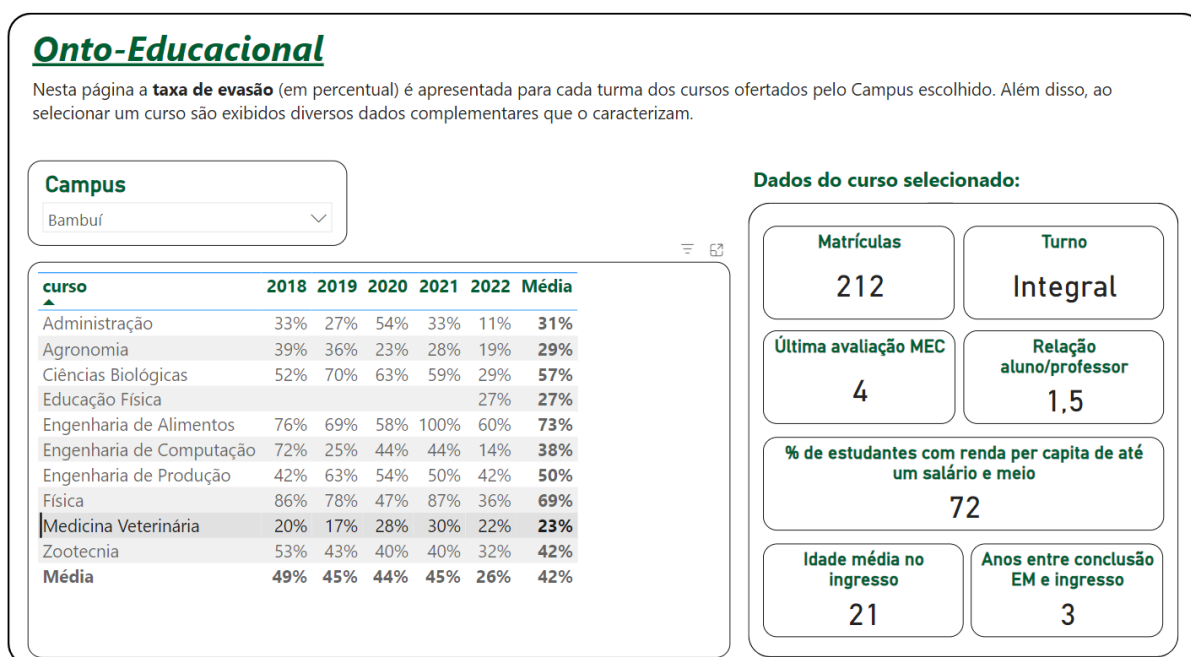
Campus	Bolsistas		Não bolsistas	
	Taxa de evasão	CR médio	Taxa de evasão	CR médio
Arcos	29,6%	50,21	68,3%	33,50
Bambuí	33,3%	64,14	44,4%	48,97
Betim	21,9%	57,15	35,0%	40,42
Congonhas	34,7%	59,19	52,2%	37,70
Formiga	29,6%	52,93	52,4%	37,37
Governador Valadares	29,3%	53,83	55,3%	38,32
Ibirité	54,5%	25,80	37,7%	31,24
Ipatinga	38,2%	44,45	56,2%	29,38
Itabirito	20,0%	50,66	20,7%	36,48
Ouro Branco	23,7%	62,93	45,3%	40,42
Ouro Preto	30,7%	53,60	64,8%	27,99
Piumhi	20,9%	53,26	51,6%	32,38
Ponte Nova	9,1%	65,01	40,4%	42,25
Ribeirão das Neves	18,7%	64,38	43,7%	38,34
Sabará	24,8%	60,04	40,2%	46,48
Santa Luzia	26,4%	59,46	44,1%	45,93
São João Evangelista	20,3%	60,22	46,1%	38,11

Fonte: Elaborada pelo autor (2024)

A quarta aba do *dashboard* permite ao usuário conhecer a taxa de evasão média de cada campus, além de segmentar este resultado pelos cursos ofertados e turmas abertas dentro do período da pesquisa (primeiro semestre de 2018 ao segundo semestre de 2022). Optou-se pela apresentação dos resultados no formato de uma tabela por campus, pois caso todos os cursos fossem disponibilizados em uma mesma estrutura, a visualização ficaria prejudicada em função da quantidade de linhas necessárias para cobrir todos os registros. No *dashboard* também são exibidos atributos adicionais para cada curso selecionado: número de matrículas, turno, última avaliação do MEC, relação aluno/professor,

percentual de estudantes com renda *per capita* de até um salário mínimo e meio, idade média no ingresso e quantidade de anos entre a conclusão do Ensino Médio e o ingresso no curso. Na Figura 56, a quarta aba do *dashboard* traz os dados dos cursos do Campus Bambuí, tendo o curso de Medicina Veterinária selecionado.

Figura 56: Quarta aba do *dashboard*



Fonte: Elaborada pelo autor (2024)

Seis *campi* do IFMG oferecem apenas um curso de graduação em suas unidades: Arcos, Ibitiré, Ipatinga, Itabirito, Piumhi e Ponte Nova. Optou-se por apresentar conjuntamente os dados desses *campi* na Tabela 26, otimizando o espaço do texto e as considerações específicas sobre cada um deles. É válido mencionar que as letras utilizadas em frente aos nomes dos cursos na tabela de cada campus referem-se à seguinte padronização: (B) Bacharelado; (L) Licenciatura; (T) Tecnológico.

No Campus Arcos observa-se que o percentual de evasão nas turmas é alto, fazendo com que a média geral do campus seja de 60%. A turma que ingressou em 2021 durante a pandemia de COVID-19 obteve a taxa mais alta, porém mesmo o resultado da turma mais recente (2022) é preocupante, pois quase metade dos ingressantes já havia evadido no final do primeiro ano. O contexto oposto é encontrado no curso de Bacharelado em Engenharia Elétrica do Campus Itabirito, cuja taxa média de evasão é baixa quando comparada com o mesmo curso ofertado no Campus Ipatinga. Há de se destacar, ainda, o único curso tecnológico da listagem, ofertado no Campus Ponte Nova, que obteve 35% de taxa média de evasão considerando as três turmas disponíveis.

Tabela 26: Evasão por curso e turmas nos *campi* que ofertam apenas um curso

Campus	Curso	2018	2019	2020	2021	2022	Média
Arcos	Engenharia Mecânica (B)	75%	62%	45%	78%	42%	60%
Ibirité	Engenharia de Controle e Automação (B)		65%	31%	36%	18%	38%
Ipatinga	Engenharia Elétrica (B)	59%	63%	48%	60%	33%	53%
Itabirito	Engenharia Elétrica (B)		35%	26%	19%	7%	22%
Piumhi	Engenharia Civil (B)	48%	45%	34%	73%	31%	46%
Ponte Nova	Processos Gerenciais (T)			37%	41%	26%	35%

Fonte: Elaborada pelo autor (2024)

O Campus Bambuí possui a maior quantidade de cursos avaliados, conforme apresentado na Tabela 27. É possível verificar grande variação no comportamento da evasão entre os cursos e, até mesmo, entre as turmas de um mesmo curso. A situação dos cursos de Bacharelado em Engenharia de Alimentos, Licenciatura em Física e Licenciatura em Ciências Biológicas é preocupante, com taxa de evasão média acima de 50%, chegando a 73% no caso do primeiro curso citado. Destaque positivo para três cursos de Bacharelado (Medicina Veterinária, Agronomia e Administração) com as menores taxas médias, excluindo neste caso o curso de Licenciatura em Educação Física por ter apenas uma turma.

Tabela 27: Evasão por curso e turmas no Campus Bambuí

Curso	2018	2019	2020	2021	2022	Média
Administração (B)	33%	27%	54%	33%	11%	32%
Agronomia (B)	39%	36%	23%	28%	19%	29%
Ciências Biológicas (L)	52%	70%	63%	59%	29%	55%
Educação Física (L)					27%	27%
Engenharia de Alimentos (B)	76%	69%	58%	100%	60%	73%
Engenharia de Computação (B)	72%	25%	44%	44%	14%	40%
Engenharia de Produção (B)	42%	63%	54%	50%	42%	50%
Física (L)	86%	78%	47%	87%	36%	67%
Medicina Veterinária (B)	20%	17%	28%	30%	22%	23%
Zootecnia (B)	53%	43%	40%	40%	32%	42%

Fonte: Elaborada pelo autor (2024)

A evasão média nos dois cursos ofertados pelo Campus Betim está próxima, com destaque para o Bacharelado em Engenharia de Controle e Automação, curso no qual a taxa vem sendo reduzida desde a turma de 2019 (Tabela 28).

Tabela 28: Evasão por curso e turmas no Campus Betim

Curso	2018	2019	2020	2021	2022	Média
Engenharia de Controle e Automação (B)	40%	42%	29%	26%	14%	30%
Engenharia Mecânica (B)	53%	34%	45%	27%	32%	38%

Fonte: Elaborada pelo autor (2024)

No Campus Congonhas dois cursos possuem taxa de evasão média superior a 50%, sendo um Bacharelado e uma Licenciatura (Tabela 29). A evasão média no curso de Bacharelado em Engenharia Mecânica chegou a 67%, resultado muito superior ao obtido pelo curso de mesma nomenclatura ofertado pelo Campus Betim, porém próximo ao do Campus Arcos. Uma das ações que a instituição poderia adotar neste caso de variação no resultado de um mesmo curso ofertado por vários *campi* seria a identificação de pontos fortes daquele com a menor taxa de evasão, propondo a adequação desses fatores nos cursos com taxas mais elevadas.

Tabela 29: Evasão por curso e turmas no Campus Congonhas

Curso	2018	2019	2020	2021	2022	Média
Engenharia de Produção (B)	54%	53%	52%	28%	27%	43%
Engenharia Mecânica (B)	62%	71%	67%	67%		67%
Física (L)	66%	49%	67%	65%	63%	62%
Letras (L)		32%	50%	36%	34%	38%

Fonte: Elaborada pelo autor (2024)

Na Tabela 30 são expostos os dados do Campus Formiga, onde cinco cursos são ofertados, sendo três bacharelados, uma licenciatura e um tecnológico. O curso de Tecnologia em Gestão Financeira obteve a segunda mais baixa média de evasão do campus, contudo desde 2021 não foram abertas novas turmas. Destaca-se o Bacharelado em Ciência da Computação com taxa média de 38%, mesmo em um curso que tradicionalmente apresenta elevados índices de desistência. O resultado mais preocupante é

do curso de Licenciatura em Matemática, com média acima de 50% em todas as turmas analisadas.

Tabela 30: Evasão por curso e turmas no Campus Formiga

Curso	2018	2019	2020	2021	2022	Média
Administração (B)	65%	56%	48%	58%	33%	52%
Ciência da Computação (B)	54%	41%	29%	32%	34%	38%
Engenharia Elétrica (B)	47%	64%	46%	42%	27%	46%
Gestão Financeira (T)	54%	39%	28%			40%
Matemática (L)	62%	61%	55%	68%	55%	60%

Fonte: Elaborada pelo autor (2024)

No Campus Governador Valadares são ofertados três cursos de bacharelado na área das Engenharias, além de um tecnológico (Tabela 31). O Bacharelado em Engenharia Civil obteve o melhor resultado, com taxa média de 27% de evasão, enquanto os outros dois cursos de engenharia tiveram índices a partir de 46%, indicando a necessidade de uma avaliação institucional para promover melhorias com vistas a nivelar os resultados desta área. O curso de Tecnologia em Gestão Ambiental está com a maior taxa média de evasão do Campus, embora o indicador tenha melhorado desde a oferta da primeira turma.

Tabela 31: Evasão por curso e turmas no Campus Governador Valadares

Curso	2018	2019	2020	2021	2022	Média
Engenharia Ambiental e Sanitária (B)	61%	57%	63%	35%	63%	56%
Engenharia Civil (B)		40%	32%	20%	14%	27%
Engenharia de Produção (B)	54%	49%	41%	55%	32%	46%
Gestão Ambiental (T)	88%	78%	60%	56%	22%	61%

Fonte: Elaborada pelo autor (2024)

Dentre os cursos do Campus Ouro Branco existe grande variação nas taxas de evasão, de acordo com a Tabela 32. Enquanto o Bacharelado em Administração e a Licenciatura em Pedagogia têm 28% e 29% de evasão média, respectivamente, a taxa do curso de Bacharelado em Engenharia Metalúrgica é superior ao dobro deste índice, requerendo atenção da instituição para promover ações que busquem identificar os motivos de tal ocorrência.

Tabela 32: Evasão por curso e turmas no Campus Ouro Branco

Curso	2018	2019	2020	2021	2022	Média
Administração (B)	33%	38%	28%	28%	13%	28%
Engenharia Metalúrgica (B)	79%	77%	52%	63%	26%	59%
Pedagogia (L)	34%	29%	30%	35%	18%	29%
Sistemas de Informação (B)	67%	61%	48%	36%	16%	46%

Fonte: Elaborada pelo autor (2024)

A Tabela 33 retorna os dados do Campus Ouro Preto, onde acontece a oferta de cinco cursos entre licenciaturas e tecnológicos. Fica evidente que o curso de Licenciatura em Física precisa de maior atenção por parte dos gestores acadêmicos do IFMG, pois a taxa de evasão média nas cinco turmas analisadas é igual a 77%, um índice muito elevado que pode, eventualmente, inviabilizar a oferta de novas entradas. Com exceção do curso de Tecnologia em Gastronomia, sem abertura de turmas desde 2021, os outros cursos também requerem ponderações institucionais sobre as causas da evasão e quais estratégias poderiam ser aplicadas para que os resultados melhorem.

Tabela 33: Evasão por curso e turmas no Campus Ouro Preto

Curso	2018	2019	2020	2021	2022	Média
Conservação e Restauro (T)	55%	63%	66%	48%	28%	52%
Física (L)	83%	82%	72%	81%	67%	77%
Gastronomia (T)	52%	32%	47%			44%
Geografia (L)	62%	71%	65%	53%	43%	59%
Gestão da Qualidade (T)	56%	53%	50%	59%	41%	52%

Fonte: Elaborada pelo autor (2024)

No Campus Ribeirão das Neves os dois cursos de graduação ofertados foram analisados, sendo que os resultados dispostos na Tabela 34 indicam certa variação entre a taxa de evasão média do Bacharelado em Administração e da Tecnologia em Processos Gerenciais, a despeito de ambos estarem relacionados com a mesma área de formação.

Tabela 34: Evasão por curso e turmas no Campus Ribeirão das Neves

Curso	2018	2019	2020	2021	2022	Média
Administração (B)	34%	33%	17%	37%	22%	29%
Processos Gerenciais (T)	45%	48%	58%	70%	25%	49%

Fonte: Elaborada pelo autor (2024)

A situação encontrada no Campus Ribeirão das Neves é similar à registrada no Campus Sabará, pois os cursos de Bacharelado em Administração e Tecnologia em Processos Gerenciais fazem parte dos que foram analisados (Tabela 35). A diferença, contudo, é que em Sabará o primeiro curso teve sua oferta iniciada quando do encerramento do segundo. Tal fato não se refletiu em mudanças nas taxas de evasão, pois a média final dos dois cursos difere em somente um ponto percentual. Outros dois bacharelados são mantidos no Campus, sendo que o de Sistemas de Informação tem taxa média de evasão superior a 50%. O curso de Tecnologia em Logística melhorou suas taxas nas duas últimas turmas em comparação com as anteriores.

Tabela 35: Evasão por curso e turmas no Campus Sabará

Curso	2018	2019	2020	2021	2022	Média
Administração (B)			30%	40%	34%	35%
Engenharia de Controle e Automação (B)		31%	53%	46%	35%	41%
Logística (T)	50%	50%	53%	26%	15%	39%
Processos Gerenciais (T)	32%	40%				36%
Sistemas de Informação (B)	44%	37%	45%	43%	21%	54%

Fonte: Elaborada pelo autor (2024)

Os cursos do Campus Santa Luzia pertencem a áreas similares de formação, motivo que talvez justifique a proximidade das taxas médias de evasão encontradas na Tabela 36. Seria interessante observar institucionalmente as ações diferenciadas que vêm sendo executadas pelo curso de Bacharelado em Arquitetura e Urbanismo para, em seguida, avaliar a possibilidade de adequação nos outros cursos.

Tabela 36: Evasão por curso e turmas no Campus Santa Luzia

Curso	2018	2019	2020	2021	2022	Média
Arquitetura e Urbanismo (B)	39%	34%	34%	40%	34%	36%
Design de Interiores (T)	35%	44%	38%	49%	31%	39%
Engenharia Civil (B)	69%	42%	34%	52%	34%	46%

Fonte: Elaborada pelo autor (2024)

O último campus investigado fica em São João Evangelista e possui seis cursos no seu catálogo, quatro bacharelados e duas licenciaturas. Dois deles se destacam com as taxas de evasão mais baixas da unidade, Bacharelado em Administração (26%) e Bacharelado em Agronomia (29%). Por outro lado, o curso com a taxa mais alta é a Licenciatura em Matemática, seguida com proximidade pelos bacharelados em Engenharia Florestal e Sistemas de Informação.

Tabela 37: Evasão por curso e turmas no Campus São João Evangelista

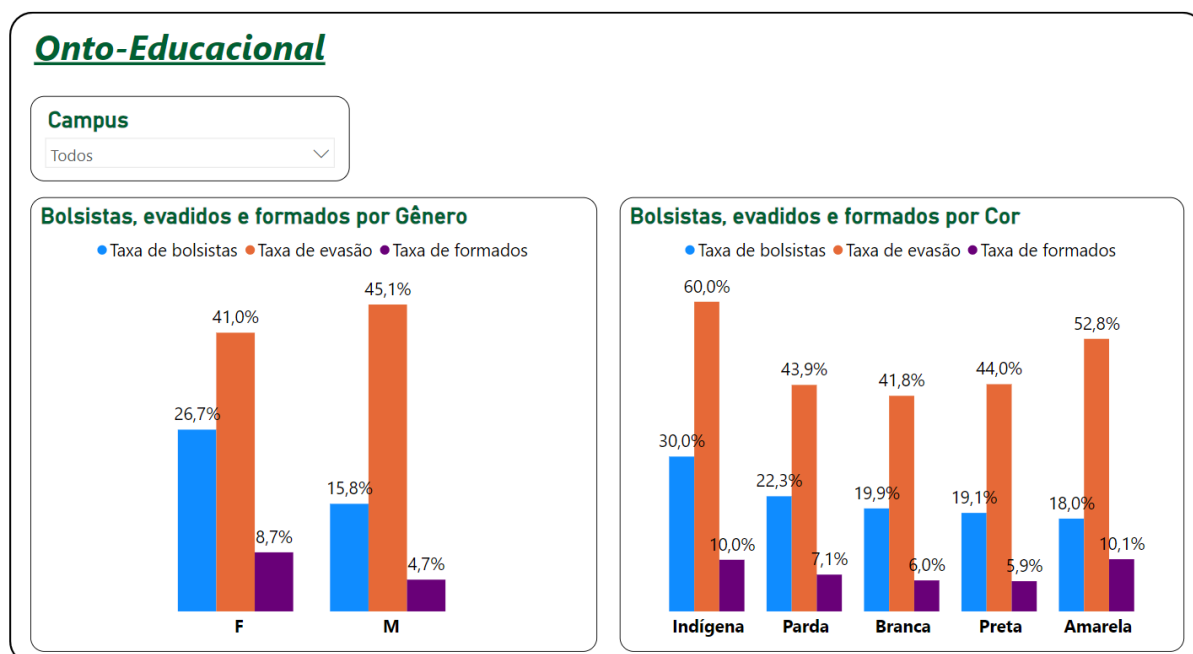
Curso	2018	2019	2020	2021	2022	Média
Administração (B)	28%	35%	30%	23%	11%	25%
Agronomia (B)	21%	36%	41%	37%	11%	29%
Ciências Biológicas (L)		38%	50%	44%	6%	35%
Engenharia Florestal (B)	56%	50%	53%	51%	15%	45%
Matemática (L)	46%	47%	51%	45%		47%
Sistemas de Informação (B)	62%	54%	46%	39%	23%	45%

Fonte: Elaborada pelo autor (2024)

A aba seguinte do *dashboard* exibe, para cada campus, o percentual de bolsistas, a taxa de evasão e o percentual de estudantes formados por gênero e cor. Na Figura 57 constam os resultados gerais do IFMG. Quanto ao gênero, observa-se que dentre os estudantes identificados pelo gênero Feminino, 26,7% são bolsistas, 41% evadiram e 8,7% se formaram. No caso do gênero Masculino, a proporção de bolsistas é de 15,8%, a taxa de evasão é igual a 45,1% e o percentual de formandos alcança 4,7%. Ressalta-se a diferença na proporção de formandos quando comparados os dois gêneros, com o resultado do Feminino sendo quase o dobro do Masculino. No caso da segmentação por cor, existe uma proximidade entre os percentuais dos três indicadores quando consideradas as opções

parda, branca e preta. As opções indígena e amarela requerem uma análise à parte, dada a pequena quantidade de matrículas existentes no conjunto de dados utilizado.

Figura 57: Quinta aba do *dashboard*



Fonte: Elaborada pelo autor (2024)

Considerando a maior variação nos resultados obtida na segmentação por gênero, a Tabela 38 visa relacionar os três indicadores por campus, permitindo identificar eventuais diferenças entre unidades com mesmo porte, área geográfica, similaridade de cursos ofertados, dentre outras possibilidades de agrupamento.

Tabela 38: Proporção de bolsistas, evadidos e formados por campus

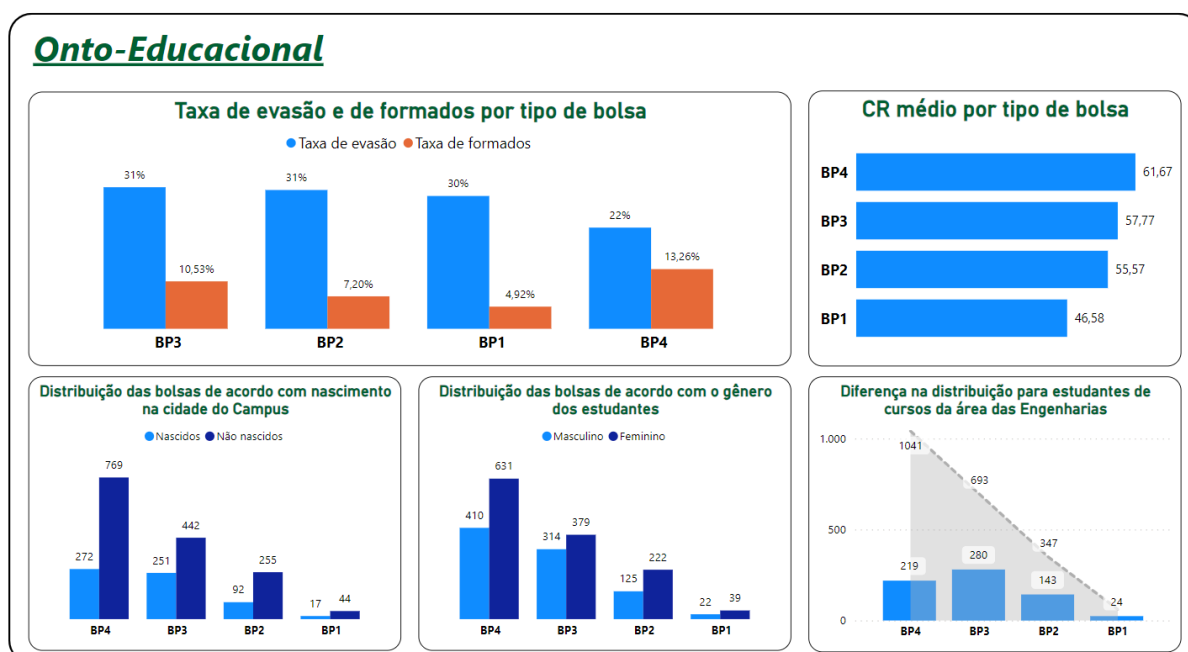
Campus	Feminino			Masculino		
	Bolsistas	Evadidos	Formados	Bolsistas	Evadidos	Formados
Arcos	28,6%	67,9%	3,6%	11,9%	61,9%	0,6%
Bambuí	26,5%	41%	4%	18,8%	42,9%	4,2%
Betim	14,9%	37,8%	0%	6,2%	33,1%	0,6%
Congonhas	26,1%	49,3%	2,9%	14,1%	48,4%	2,9%
Formiga	26,5%	46,4%	20,1%	21,3%	47,5%	7,6%
Governador Valadares	29,2%	48,5%	3,9%	18,4%	49,9%	4%
Ibirité	15%	45%	0%	7,6%	38,1%	0%
Ipatinga	35,1%	59,5%	0%	16,5%	50,4%	0%
Itabirito	29%	32,3%	0%	11,9%	17,9%	0,7%
Ouro Branco	31,2%	32%	15,2%	13,1%	50,3%	2%
Ouro Preto	19%	55,4%	4,3%	10,7%	64,8%	1,4%

Piumhi	28,1%	46,1%	1,1%	16,2%	44,1%	0%
Ponte Nova	23%	32,8%	14,8%	14,5%	36,4%	18,2%
Ribeirão das Neves	25,4%	38,2%	19,3%	14,3%	39,1%	18,9%
Sabar	18,1%	31,3%	16%	10,8%	41,2%	9,6%
Santa Luzia	16,8%	36,3%	8,3%	15,2%	48,8%	0,8%
So Joo Evangelista	39,6%	35,2%	5,3%	28,5%	39,6%	3,8%

Fonte: Elaborada pelo autor (2024)

Na sexta aba do *dashboard* buscou-se segmentar os dados por tipo de bolsa distribuda aos estudantes. Os grficos dispostos na Figura 58 trazem informaes teis para compreender a relao de cada tipo de auxlio distribudo com outras variveis. Por exemplo, quando a taxa de evaso  analisada em conjunto com os tipos de bolsa, tem-se grande proximidade entre BP1, BP2 e BP3, indicando que cerca de um tero dos estudantes que receberam essas bolsas evadiram de seus cursos. Quanto ao percentual de formados, os estudantes vinculados a BP1 (bolsa de maior valor) tiveram o resultado mais baixo; em contrapartida, o resultado mais alto foi obtido por aqueles recebedores de BP4, indicando que no somente o valor do auxlio  capaz de impulsionar este indicador. Situao similar acontece ao verificar o coeficiente de rendimento mdio por tipo de bolsa, sugerindo que esta varivel possui importncia para identificao se o estudante est mais propenso a se formar.

Figura 58: Sexta aba do *dashboard*

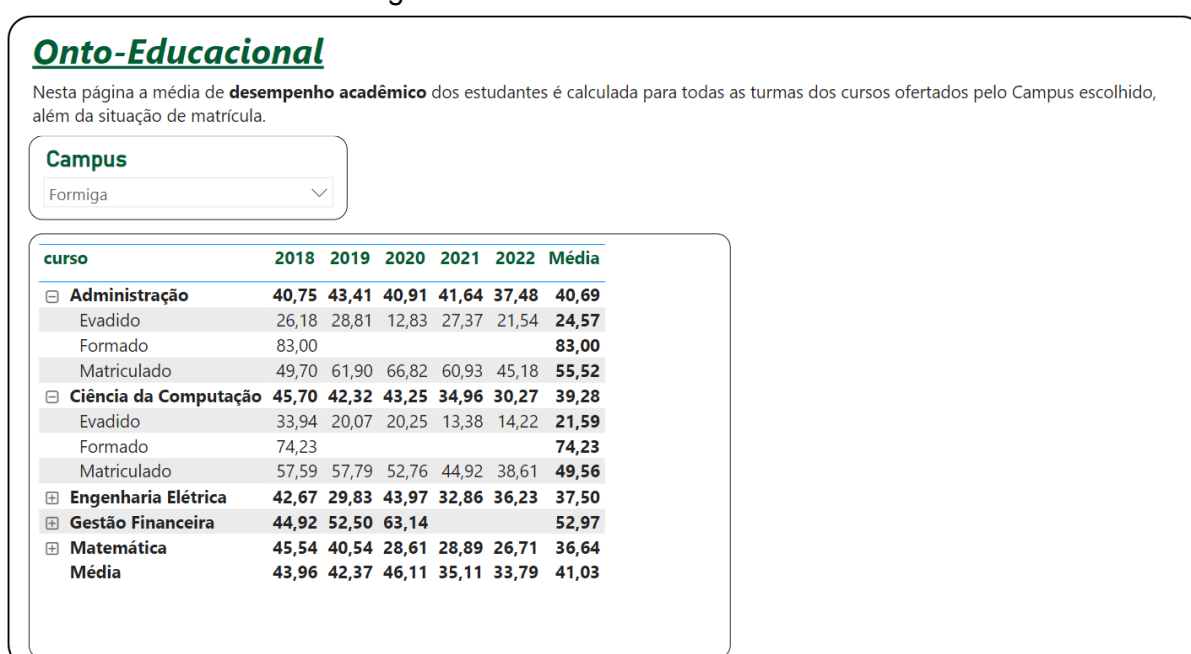


Fonte: Elaborada pelo autor (2024)

A stima aba do *dashboard* permite que o usurio analise a mdia do desempenho acadmico dos estudantes de cada curso de um campus, segmentando esses dados por

turma e por situação de matrícula (evadido, formado e matriculado). Devido ao grande volume de dados, optou-se por não reproduzir no corpo da tese os resultados apresentados nesta aba. A título de exemplo, a Figura 59 exibe os dados do Campus Formiga. É possível verificar certa diferença entre o desempenho dos estudantes evadidos para aqueles que ainda estão matriculados no curso, inclusive dentro de uma mesma turma. Percebe-se, também, diferença maior ainda quando o desempenho dos estudantes formados é analisado, pois a média do coeficiente de rendimento acadêmico é notadamente superior à obtida pelos evadidos e matriculados.

Figura 59: Sétima aba do *dashboard*



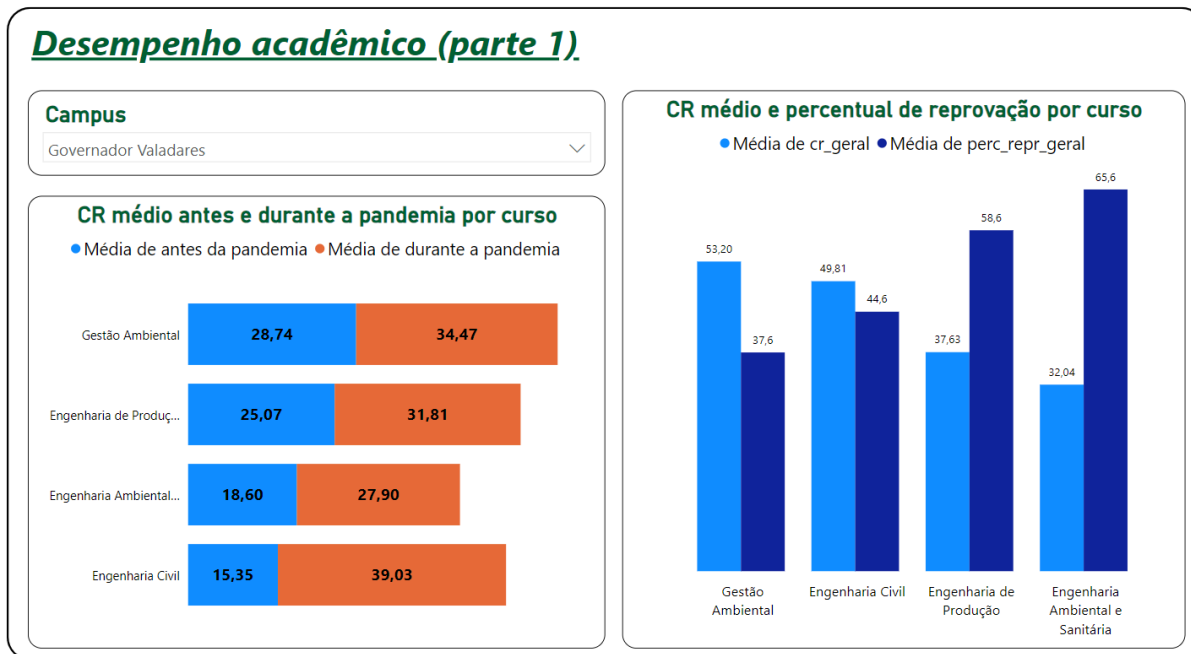
Fonte: Elaborada pelo autor (2024)

O desempenho acadêmico faz parte do contexto dos dados explorados no formato de gráficos na oitava e nona abas do *dashboard*. Conforme verifica-se na Figura 60, a primeira parte objetiva retornar, para cada curso do campus escolhido, o coeficiente de rendimento médio antes e durante a pandemia, além da média desse indicador em conjunto com a média do percentual geral de reprovações. As turmas de 2018 e 2019 correspondem ao período nomeado como antes da pandemia, enquanto as turmas de 2020, 2021 e 2022 ao período durante a pandemia.

Utilizando o Campus Governador Valadares como exemplo, observa-se incremento na média do desempenho de todos os cursos na comparação dos períodos antes e durante a pandemia, podendo servir de ponto de partida para uma reflexão dos gestores institucionais sobre como as atividades acadêmicas foram conduzidas em ambos os momentos. Quanto ao percentual de reprovação por curso, chama atenção certa relação

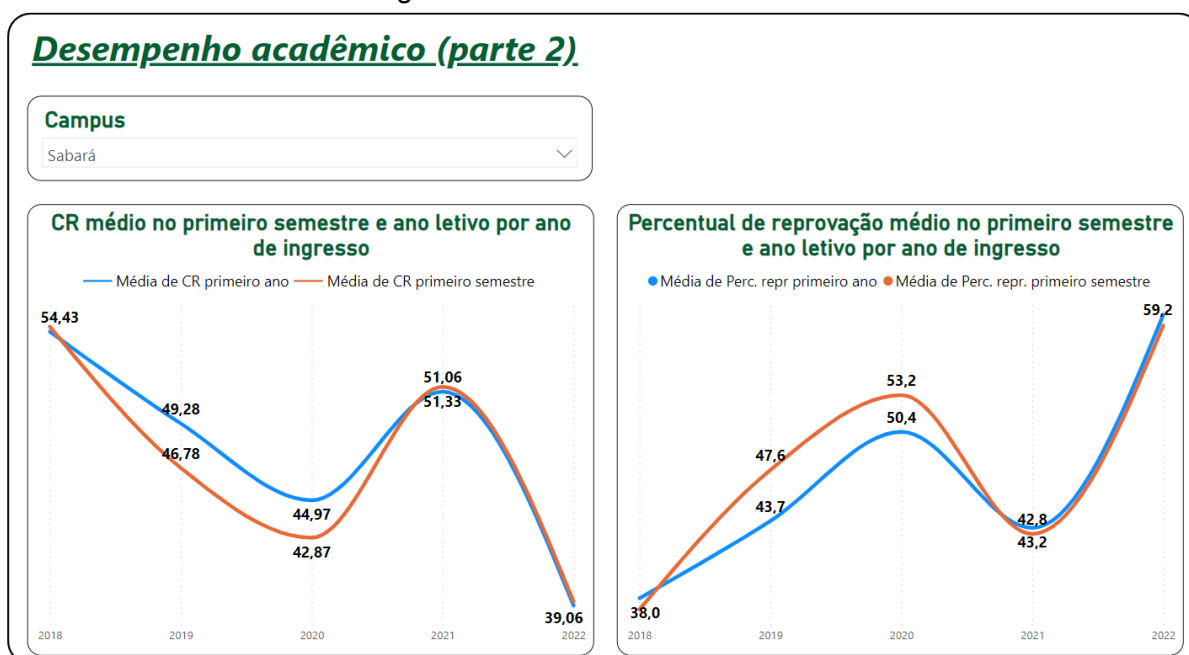
proporcional entre este indicador e o coeficiente de rendimento: quando o desempenho acadêmico é alto, o percentual de reprovações dos estudantes tende a diminuir.

Figura 60: Oitava aba do *dashboard*



Fonte: Elaborada pelo autor (2024)

A última aba traz a segunda parte das análises gráficas sobre o desempenho acadêmico, exibindo a média do coeficiente de rendimento no primeiro ano e no primeiro semestre letivo do campus escolhido dentro do escopo da pesquisa (2018 a 2022), além da média do percentual de reprovação nesses mesmos moldes. A Figura 61 contém dois gráficos gerados com os dados do Campus Sabará, permitindo novamente reconhecer a relação inversamente proporcional entre o coeficiente de rendimento e o percentual de reprovação; quando o CR está elevado, o percentual de reprovação é baixo e vice-versa. Interessante notar, ainda, as alterações nas variáveis quando os anos são comparados. Neste exemplo, enquanto nos anos de 2018, 2021 e 2022 as duas variáveis de coeficiente de rendimento possuem médias próximas, em 2019 e 2020 os resultados agregados no primeiro ano letivo foram superiores aos obtidos ao final do primeiro semestre letivo.

Figura 61: Nona aba do *dashboard*

Fonte: Elaborada pelo autor (2024)

Encerrada a construção do protótipo de *dashboard* gerencial, passou-se à análise dos dados para compreender as relações entre os construtos. Tal atividade contou com a participação direta das duas especialistas de domínio que haviam atuado na captura e formalização da ontologia, ambas servidoras do IFMG, uma na área de ensino (denominada “Especialista A”) e a outra na área de assistência estudantil (“Especialista B”). Optou-se pela realização de um grupo focal *online* entre o pesquisador e as especialistas para que as impressões individuais pudessem ser compartilhadas e novos conhecimentos fossem construídos coletivamente. A escolha das especialistas foi intencional, dado o encontro entre a experiência de ambas nas suas áreas de atuação e o tema central desta tese.

A preparação do grupo focal envolveu o convite às especialistas e, após os devidos aceites, a disponibilização da versão final da Onto-Educacional, das questões de competência e suas respostas, além do endereço virtual de acesso ao *dashboard* gerencial para que elas realizassem uma análise prévia dos dados antes da reunião conjunta. Somente uma sessão do grupo focal ocorreu na plataforma *Google Meet*, teve duração de aproximadamente 70 minutos e foi conduzida pelo pesquisador. O roteiro da reunião dispunha do tópico “análise das relações entre a evasão escolar, desempenho acadêmico e distribuição de auxílio socioeconômico”.

Iniciada a reunião, o pesquisador agradeceu as especialistas pela presença e discorreu, brevemente, acerca das atividades conduzidas após a presença delas na concepção da ontologia, culminando com os resultados proporcionados pelo ambiente informacional construído sob a estrutura modelada na Onto-Educacional.

Na sequência, para que as discussões tivessem início, passou a palavra para a Especialista A, que afirmou ser “um sonho conseguir agrupar e visualizar esses dados”, pois muitas vezes as decisões da gestão acadêmica são tomadas com base em achismos. Ela continuou dizendo que, embora haja um esforço para inserir os dados institucionais em plataformas governamentais, como a Nilo Peçanha, nem sempre é possível cruzar esses dados com outros. Ainda na visão desta especialista, a análise de dados integrados faz muita diferença para suscitar questões a serem discutidas em conjunto com possibilidades de intervenção por parte da instituição. Finalizando sua primeira participação, ela conclui que “o acesso a uma plataforma que agrupa os dados educacionais é fundamental para o trabalho da gestão”.

A Especialista B ressaltou, em sua primeira fala, a validade do estudo (no seu entendimento) por ter abrangido todos os *campi* do IFMG, pois não “é uma situação específica de alguma unidade ou curso, mas sim que os auxílios realmente impactam no cotidiano dos estudantes em toda a instituição”. Ela percebe, ainda, que os auxílios tiveram importante papel durante a pandemia de COVID-19, contribuindo de diferentes formas para que os estudantes mantivessem seu vínculo com a instituição, seja utilizando o recurso para comprar planos de internet, pagando despesas da residência ou adquirindo itens de alimentação.

Tratando sobre a análise das relações envolvendo os três construtos, a Especialista A mencionou que sua impressão geral apontava para diferença na evasão e no desempenho acadêmico entre bolsistas e não bolsistas, “mas não tinha noção que ela fosse tão considerável como os resultados mostraram”. Especificamente sobre o desempenho acadêmico, a mesma Especialista afirma que, considerando sua observação da realidade no campus, o fato de receber uma bolsa gera mais comprometimento com o curso e com a instituição, pois os estudantes apoiados vão se esforçar para continuar recebendo o auxílio. Muitos deles “sabem que, sozinhos ou mesmo com a ajuda da família, não conseguem se manter estudando”.

A Especialista B indicou percepção semelhante ao afirmar que o bolsista parece ter mais dedicação nas suas tarefas escolares para não perder o direito ao benefício, o que se traduz como impacto positivo no desempenho acadêmico. Para ela, tem-se “uma cadeia de acontecimentos interligados, ele recebe a bolsa, estuda e se dedica mais, tem melhores notas e se forma”. Os estudantes bolsistas, segundo ela, “têm uma consciência muito grande do que querem, vieram para o IFMG para estudar e vão fazer de tudo para continuar até o final, pois dão grande valor naquilo que conseguiram.”.

Quando discutidos os resultados da evasão e sua relação com a distribuição do auxílio socioeconômico, a Especialista A entende que “o comprometimento assumido para concluir o curso, principalmente pelo bolsista, é a base para entender a diferença da evasão

entre os bolsistas e não bolsistas, pois eles precisam daquele suporte para continuar a perseguir este objetivo”. A Especialista B menciona que “o percentual de evasão entre os bolsistas está perto de um número que entendo ser aceitável para cursos superiores, pois são muitos fatores que levam os alunos a evadirem e é difícil para a instituição controlar todos. Já a proporção de evadidos entre os não bolsistas é muito alta, de quase 50%, o que poderia até mesmo inviabilizar a oferta de algum curso se ela for constante em várias turmas”.

Ambas as especialistas consentiram que o auxílio direcionado para aqueles estudantes que não têm origem na cidade do campus onde o curso é mantido possui especial impacto na evasão, dada inexistência de uma rede de apoio que contribua com alimentação, moradia e transporte para o local de estudo. A Especialista B nota, em seu cotidiano, que aqueles alunos que vêm de outros estados e cidades demonstram maior interesse nas atividades acadêmicas e interação com a comunidade, “pois entendem que estão lá para estudar e então se dedicam mais a isso”.

Outro aspecto trazido pela Especialista A denota a importância dos auxílios para o controle da evasão. Segundo ela, alguns cursos acontecem em período integral, fazendo com que os alunos tenham que estudar o dia todo e não consigam aumentar a renda familiar por meio do trabalho. Essa situação é muito diferente no contexto dos cursos realizados no período noturno, quando os estudantes “conseguem dar continuidade nos seus cursos por causa da renda obtida em trabalhos realizados durante o dia”. A Especialista B complementa que as particularidades de cada cidade onde o IFMG possui campus contribuem, de certa forma, para a demanda por auxílio socioeconômico: “Alguns *campi* contam com muitos estudantes que vêm de outras cidades para morar, em outros os estudantes precisam contratar serviço de van para chegar na escola, na região metropolitana de Belo Horizonte o custo de alimentação costuma ser mais caro... Acho que essa é uma das vantagens da bolsa permanência ser paga em dinheiro, porque assim cada estudante pode utilizá-la de acordo com a sua necessidade”.

Nesse momento, a Especialista B tratou sobre o papel da instituição quanto ao que pode ser feito para melhorar os indicadores de evasão. Na sua visão, gestores acadêmicos e professores precisam repensar aspectos da oferta de cursos com evasão muito alta, objetivando entender o que está contribuindo para a saída dos estudantes e, a partir desta análise, conceber planos que melhorem a situação como um todo. Ela ainda complementa: “Não digo que o curso precisa ser encerrado, mas que todos devem pensar juntos ações que poderiam melhorá-lo, como mudar o turno de oferta de integral para noturno ou o tempo mínimo para integralização, por exemplo”.

A Especialista A trouxe para a discussão sobre fatores relacionados com a evasão uma questão que, a seu ver, muitas vezes é negligenciada pelas instituições educacionais: a

saúde mental dos estudantes. Ela afirma que notou, enquanto Diretora de Ensino de um campus, considerável aumento nos quadros de ansiedade entre os estudantes, com maior incidência após o início da pandemia de COVID-19 e em indivíduos do gênero feminino. Muitos estudantes ficam ansiosos pela obtenção de desempenho acadêmico satisfatório para continuidade do curso, outros por medo de determinadas disciplinas e para a realização do trabalho de conclusão de curso, enquanto a cobrança externa de familiares e da própria sociedade são outros fatores observados como causas de quadros de ansiedade. Ela ressalta que, embora alguns *campi* ofereçam algum tipo de suporte psicológico para os estudantes, a demanda fica muito acima das capacidades de atendimento institucional.

As duas especialistas mencionaram o baixo nível de conhecimentos prévios em disciplinas do Ensino Médio e a falta de informações mínimas sobre o curso como outros fatores que impactam na evasão do IFMG. A Especialista B, que atua no setor de assistência estudantil, ressalta que alguns candidatos aprovados via SISU sequer “sabem a localização da cidade onde irão morar e estudar”. Ela complementa: “Os jovens que entram hoje no ensino superior não têm a mesma percepção da importância do estudo para a vida profissional e pessoal que as gerações anteriores tinham. Muitos nem conhecem direito o curso que estão fazendo. Quando chega um momento mais difícil, por causa de disciplinas ou professores, preferem abandoná-lo porque ainda não viram o tamanho da perda”.

A evasão também é impactada pela oferta de cursos superiores ofertados totalmente à distância, sem a necessidade de presença física na instituição, de acordo com a Especialista A. Ela entende que “para muitos estudantes com dificuldades financeiras, essa passou a ser uma opção mais viável do que se mudar para uma cidade desconhecida e frequentar aulas todos os dias presencialmente. Mesmo que não seja cobrada mensalidade no IFMG, o estudante tem diversos outros gastos que dificultam a continuidade do seu vínculo”. A Especialista B acrescenta que é justamente neste contexto que a oferta de auxílio socioeconômico pode ser útil, uma vez que a bolsa permanência seria “outro fator motivacional para que o aluno optasse por estudar em uma instituição com qualidade de ensino reconhecida pelo mercado, mesmo com as dificuldades de morar em outra cidade e frequentar presencialmente o curso”. Para a Especialista A, um desafio a ser enfrentado pelo IFMG consiste na construção de um modelo híbrido de ensino, com ensino à distância em parte da carga horária dos cursos, aproximando assim o ambiente acadêmico da realidade vivenciada pelos estudantes fora da instituição.

No momento em que a discussão tratou sobre o impacto de cada tipo de bolsa na evasão e no desempenho, a Especialista A afirmou que, embora o valor pago pela BP4 seja o menor, “já é um diferencial para muitos alunos vulneráveis, ajudando, por exemplo, na aquisição de planos de internet e alimentação”. Ou seja, dado um contexto de grande vulnerabilidade, o valor promove algum tipo de diferença na vida daqueles que o recebem. A

especialista exemplifica mencionando o caso daqueles estudantes que residem em repúblicas, “onde as despesas são menores por causa da divisão com outras pessoas”.

A Especialista B informa que uma das estratégias adotadas pela instituição para alcançar um número maior de beneficiários de auxílios, em função da grande demanda gerada por indivíduos com necessidades socioeconômicas, é a distribuição de bolsas com menor valor para um público mais amplo quantitativamente tratando. Ela relata que “foi bom confirmar em números que a bolsa de menor valor faz diferença para nossos bolsistas, pois tínhamos a impressão que ela ajudava a diminuir pequenas barreiras que os estudantes enfrentam na sua vida, mas esse resultado contribui para que as análises de solicitações de bolsas sejam aprimoradas. Pode ser que, em alguns casos, o estudante bolsista BP4 até consiga completar seu curso sem a bolsa, talvez com um esforço maior da família e dele mesmo, mas o auxílio não deixa de ser um diferencial para que ele continue se esforçando”.

Uma importante reflexão foi trazida pela Especialista B quando afirma que “os estudantes que recebem a BP1 (maior valor financeiro) vêm de uma família muito vulnerável, talvez por este motivo eles não tenham obtido o maior desempenho acadêmico, pois passam por tantas dificuldades que acabam impactando no seu estudo. Muitos não têm um local adequado para estudo, outros vivem em meio a conflitos familiares, enfim, mesmo recebendo uma bolsa de valor maior, isso não é suficiente para resolver todos os seus problemas”. Ela sugere, inclusive, a realização de um estudo futuro objetivando compreender se o auxílio de maior valor contribuiu na melhoria do resultado acadêmico dos bolsistas, isto é, se houve evolução mais acentuada entre esses bolsistas quando comparados com não bolsistas de perfil semelhante.

Encerradas as considerações sobre as relações entre os construtos, a Especialista A afirmou que o estudo conduzido no âmbito desta tese “é muito importante para comprovarmos a efetividade do auxílio e requerermos a sua ampliação na instituição”. A Especialista B complementa mencionando que os resultados do estudo, baseado nos dados do próprio IFMG, mostram que o orçamento institucional precisa ser revisado pelos gestores, a fim de que sejam direcionadas mais verbas para ampliar a concessão de bolsas para estudantes com vulnerabilidades socioeconômicas e, por conseguinte, ocorra incremento no desempenho acadêmico deles e redução nos números da evasão.

Ao final da reunião, o pesquisador encerrou formalmente a sessão do grupo focal e agradeceu as especialistas por terem engrandecido o estudo com variados pontos de vista sobre os dados analisados, bem como pela avaliação da Onto-Educacional enquanto usuárias externas.

4.5 Análise sobre o impacto do PNAES na instituição

A tarefa inicial desenvolvida nesta fase consistiu na obtenção do conjunto de dados a ser trabalhado pelo método *Propensity Score Matching* (PSM). Utilizando uma consulta SQL, os registros foram recuperados do repositório intermediário criado anteriormente para integrar os dados das fontes que proveram dados para este estudo, conforme exposto na Subseção 4.3.3.2. O resultado foi exportado para uma planilha no *Microsoft Excel* com o intuito de preparar os dados antes de seu processamento.

As tarefas de preparação envolveram, inicialmente a exclusão dos atributos *campus*, *curso*, *ultima_avaliacao_mec*, *relacao_aluno_professor*, todos aqueles relacionados com o coeficiente de rendimento (exceto o *cr_geral*) e com o percentual de reprovação, *qtd_periodos_cursados*, *qtd_aprovacoes* e *qtd_reprovacoes*, com o objetivo de reduzir a dimensionalidade dos dados conforme sugere o PSM, a partir da supressão de atributos não diretamente relacionados com a variável de tratamento que, no caso, indica se um estudante é bolsista ou não. A segunda tarefa de preparação transformou em binários os atributos que apresentavam somente dois possíveis valores (*genero*, *natural_cidade_campus* e *bolsista_ae*), passando seus valores para 0 ou 1. Para encerrar a preparação do conjunto de dados, aqueles atributos com mais de dois valores possíveis foram agrupados ou desmembrados: *area_conhecimento* (agrupado em *area_engenharia_exatas*, área com a maior quantidade de registros), *turno* (agrupado em *estuda_noturno*, pois a maioria dos registros são de estudantes deste turno), *grau_academico* (desmembrado em *grau_academico_bacharelado*, *grau_academico_licenciatura*, *grau_academico_tecnologico*), *forma_ingresso* (desmembrado em *forma_ingresso_sisu*, *forma_ingresso_vestibular* e *forma_ingresso_outras*) e *cor* (desmembrado em *cor_branca*, *cor_parda* e *cor_outras*).

Ao final, foi exportado um arquivo no formato CSV composto de 10.263 registros e 20 atributos, cuja estrutura está disposta no Quadro 7, inclusive com a indicação de quais atributos representam as variáveis independentes, dependentes e de tratamento do PSM.

Quadro 7 - Atributos do conjunto de dados para execução do PSM

Nº	Atributos	Tipo	Descrição
Variáveis independentes			
1	<i>grau_academico_bacharelado</i>	Numérico binário	Vínculo com curso de Bacharelado ou não
2	<i>grau_academico_licenciatura</i>	Numérico binário	Vínculo com curso de Licenciatura ou não
3	<i>grau_academico_tecnologico</i>	Numérico binário	Vínculo com curso de Tecnologia ou não
4	<i>area_engenharias_exatas</i>	Numérico binário	Estuda em curso de exatas ou não
5	<i>estuda_noturno</i>	Numérico binário	Estuda em curso noturno ou não
6	<i>carga_horaria</i>	Numérico discreto	Carga horária do curso

7	perc_rfp_ate_1salmeio_curso	Numérico discreto	Percentual de alunos do curso com renda <i>per capita</i> familiar de até um salário mínimo e meio
8	forma_ingresso_sisu	Numérico binário	Ingressou pelo SISU ou não
9	forma_ingresso_vestibular	Numérico binário	Ingressou por vestibular ou não
10	forma_ingresso_outras	Numérico binário	Ingressou por outras formas ou não
11	genero_masculino	Numérico binário	Indica se é do gênero masculino ou não
12	cor_branca	Numérico binário	Indica se é da cor branca ou não
13	cor_parda	Numérico binário	Indica se é da cor parda ou não
14	cor_outras	Numérico binário	Indica se é de outra cor ou não
15	natural_cidade_campus	Numérico binário	Nasceu na mesma cidade do Campus ou não
16	idade_ingresso	Numérico discreto	Idade no momento do ingresso no curso
17	tempo_entre_conclusao_em_ingresso	Numérico discreto	Tempo decorrido entre a conclusão do Ensino Médio e o ingresso no curso de graduação
Variável de tratamento			
18	bolsista	Numérico binário	Indica se é bolsista da assistência estudantil ou não
Variáveis dependentes			
19	evadido	Numérico binário	Indica se evadiu do curso ou não
20	cr_geral	Numérico discreto	Coefficiente de rendimento

Fonte: Elaborada pelo autor (2024)

O arquivo CSV foi importado em um *script* no software RStudio para a execução das manipulações sobre os dados. O conteúdo do *script* e o arquivo CSV estão disponíveis no repositório virtual do projeto em <https://github.com/educardosomelo/onto-educacional/tree/main/PSM>.

É válido recordar dois resultados apontados na Subseção 4.3.3.4, quando as questões de competência que guiaram a construção da Onto-Educacional foram respondidas. Em relação ao desempenho acadêmico, observou-se uma diferença de 18,84 pontos no coeficiente de rendimento médio entre os bolsistas (média de 58,99) e os não bolsistas (média de 40,15). Quanto à evasão, a proporção entre os bolsistas foi de 26,9%, enquanto para os não bolsistas alcançou 47,5%. O Teste T apontou, em ambos os casos, que as diferenças são estatisticamente significativas para $p < 0,05$.

A questão central, ao investigar o impacto da política pública oficializada como PNAES no IFMG, é verificar se existe diferença significativa nas duas variáveis dependentes conforme a variável de tratamento e, caso exista, será que essa situação se manifesta porque os estudantes bolsistas e os não bolsistas têm desempenhos divergentes e evadem por razões ligadas ao recebimento do auxílio ou isso acontece por causa das variáveis

independentes que caracterizam o conjunto de dados e, eventualmente, colocam os estudantes dos dois grupos em situação de desigualdade? Em outras palavras, é o recebimento do auxílio socioeconômico que produz essa desigualdade ou as outras características dos estudantes?

As médias das variáveis independentes nos dois grupos (bolsistas e não bolsistas) são apresentadas na Tabela 39. A verificação se as diferenças entre as médias eram estatisticamente significativas foi obtida com a aplicação de Testes T individualizados, apontando que em 13 das 17 variáveis houve significância. Desta forma, entende-se que apenas a comparação das médias gerais do desempenho acadêmico e da proporção de evasão entre os dois grupos não permite afirmar, com total certeza, que o recebimento de auxílio socioeconômico é o fator causal, pois a diferença encontrada pode estar sofrendo impacto de outras características dos estudantes.

Tabela 39 - Médias das variáveis nos dois grupos

Variáveis	Bolsista	Não bolsista	p	Significante
grau_academico_bacharelado	0,67	0,69	0,1165	Não
grau_academico_licenciatura	0,19	0,14	< 0,001 (3,304E-06)	Sim
grau_academico_tecnologico	0,13	0,15	0,0019	Sim
area_engenharias_exatas	0,44	0,54	< 0,001 (1,917E-15)	Sim
estuda_noturno	0,56	0,59	0,0055	Sim
carga_horaria	3.285	3.230	0,0010	Sim
perc_rfp_ate_1salmeio_curso	70,53	66,62	< 0,001 (2,2E-16)	Sim
forma_ingresso_sisu	0,70	0,67	0,0035	Sim
forma_ingresso_vestibular	0,23	0,24	0,2494	Não
forma_ingresso_outras	0,05	0,08	0,0005	Sim
genero_masculino	0,40	0,57	< 0,001 (2,2E-16)	Sim
cor_branca	0,36	0,38	0,0711	Não
cor_parda	0,48	0,44	0,0011	Sim
cor_outras	0,15	0,17	0,0397	Sim
natural_cidade_campus	0,29	0,28	0,3906	Não
idade_ingresso	22,43	24,15	< 0,001 (2,2E-16)	Sim
tempo_entre_conclusao_em_ingresso	4,42	6,04	< 0,001 (2,2E-16)	Sim

Fonte: Elaborada pelo autor (2024)

Para promover uma comparação mais justa entre os resultados dos bolsistas e não bolsistas, será criado um contrafactual, isto é, um grupo de controle composto por estudantes não bolsistas com a maior probabilidade de terem as mesmas características dos estudantes bolsistas (grupo de tratamento). Essa probabilidade é o escore de propensão do método PSM, cuja estimativa é calculada com a criação de um modelo logarítmico que engloba a variável de tratamento (*bolsista*) como variável dependente e as variáveis independentes que servirão como controle. Há de ressaltar que o PSM pode encontrar dificuldades na etapa de pareamento quando existe grande quantidade de características (variáveis) a serem analisadas nos grupos de controle e tratamento, motivo pelo qual é importante definir quais irão compor o modelo logarítmico de estimativas.

Sendo assim, previamente ao cálculo do escore de propensão de cada estudante, é calculada a probabilidade individual dele ser bolsista com base nas variáveis independentes, indicando aquelas com participação direta no tratamento e que deveriam estar no modelo logarítmico. Na Tabela 40 são expostas as sete variáveis independentes com significância estatística para serem preditoras da concessão de auxílio socioeconômico.

Tabela 40 - Variáveis independentes preditoras

Variáveis	Coefficiente	Erro padrão	z	p > Z	Significante
grau_academico_licenciatura	0,2443	0,0724	3,371	0,0007	Sim
area_engenharias_exatas	-0,1797	0,0535	-3,354	0,0007	Sim
estuda_noturno	-0,1497	0,0571	-3,622	0,0008	Sim
perc_rfp_ate_1salmeio_curso	0,0103	0,0016	6,147	< 0,0001	Sim
genero_masculino	-0,5767	0,0529	-10,895	< 0,0001	Sim
cor_branca	-0,2757	0,0534	-5,160	< 0,0001	Sim
tempo_entre_conclusao_em_ingresso	-0,0420	0,0047	-8,898	< 0,0001	Sim

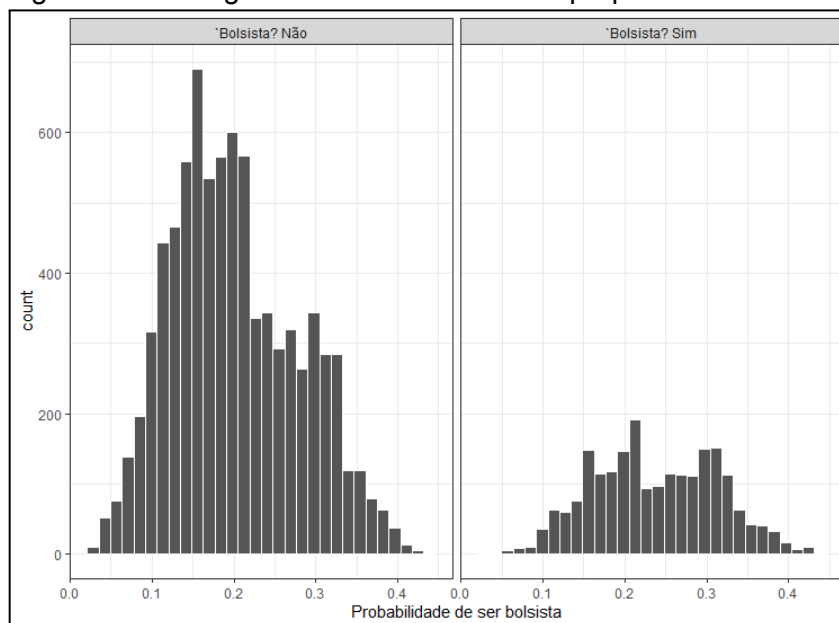
Fonte: Elaborada pelo autor (2024)

Elaborada a definição do modelo logarítmico, o escore de propensão para cada estudante foi calculado. Trata-se do valor de probabilidade predita do estudante compor o grupo de tratamento, isto é, ser bolsista, dadas as estimativas do modelo logarítmico. O escore é calculado mesmo para aqueles estudantes que fazem parte do grupo de tratamento no conjunto de dados original.

No intuito de compreender se existe balanceamento das probabilidades calculadas entre bolsistas e não bolsistas, gerou-se o histograma apresentado na Figura 62. Verifica-se uma região de suporte comum entre as probabilidades dos dois grupos. A maioria dos

valores está, em ambos os grupos, no intervalo de 0,0 a 0,4. O escore médio entre os bolsistas foi de 0,238 e de 0,201 no grupo de não bolsistas.

Figura 62 - Histograma com os escores de propensão calculados



Fonte: Elaborada pelo autor (2024)

Considerando que as probabilidades foram calculadas para todos os estudantes do conjunto de dados, executou-se a tarefa de pareamento (*matching*) visando encontrar pares com escores de propensão muito semelhantes, mas com status de tratamento diferentes. O método do vizinho mais próximo serviu de base para esta tarefa, analisando os cinco vizinhos mais próximos de cada estudante do grupo de tratamento e, em seguida, identificando no grupo de controle um indivíduo com escore de propensão igual ou aproximado. O conjunto de dados original é composto de 10.263 registros, sendo 8.121 de estudantes não bolsistas e 2.142 de bolsistas. Já o conjunto de dados pareados apresenta 4.282 registros, com igual quantidade (2.142 cada) entre bolsistas (grupo de tratamento) e não bolsistas (grupo de controle).

É importante ressaltar que o grupo de controle é composto tanto por estudantes que solicitaram auxílio quanto por aqueles que, eventualmente, não solicitaram auxílio em algum momento da sua vida acadêmica, pois este dado não foi disponibilizado pela instituição. Entretanto, como cada indivíduo do grupo de controle possui características muito próximas a um estudante do grupo de tratamento, conforme o resultado da tarefa de pareamento, a comparação entre os indivíduos se torna mais adequada, tendo como diferença principal o fato de ser bolsista ou não.

Para verificar o balanceamento entre os grupos neste conjunto de dados pareados, foi calculada a diferença das médias para as sete variáveis independentes que o compõem, seguido de um Teste T para identificar a existência de significância estatística entre as diferenças encontradas. Conforme exposto na Tabela 41, os grupos de tratamento e controle estão balanceados, não havendo diferença estatística entre as médias para todas as variáveis analisadas, fato que permite a comparação das variáveis dependentes (desempenho acadêmico e evasão) entre os grupos e a medição do impacto causado pelo tratamento.

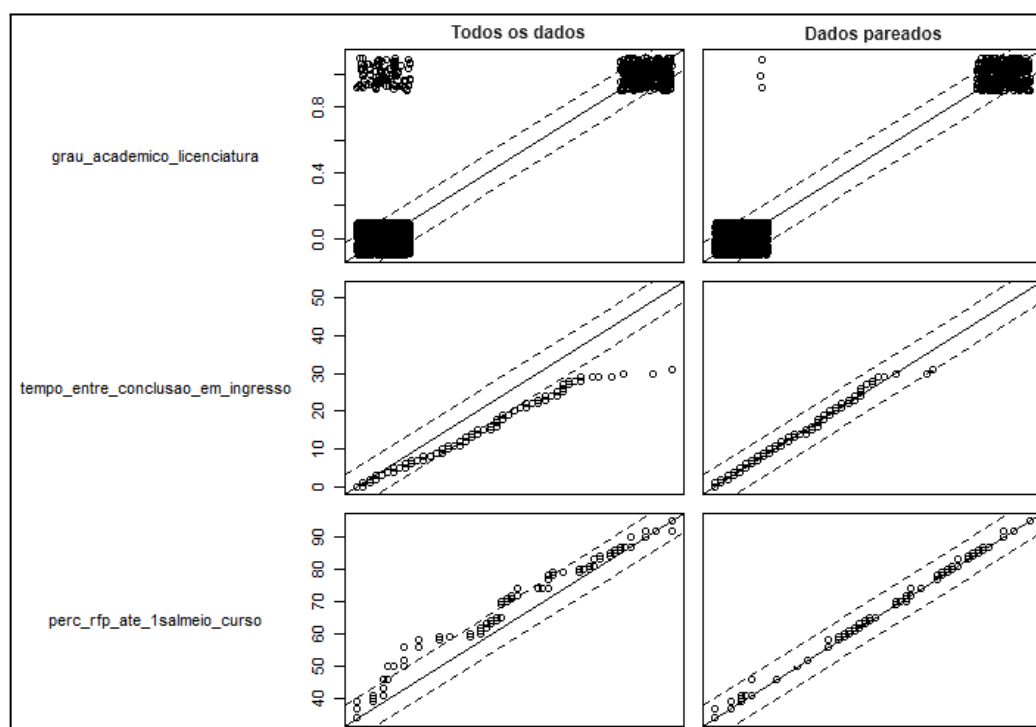
Tabela 41 - Comparação das variáveis nos grupos de tratamento e controle

Variáveis	Média do grupo de tratamento	Média do grupo de controle	Diferença média padronizada	<i>p</i>	Significante
grau_academico_licenciatura	0,19	0,18	0,0226	0,4558	Não
area_engenharias_exatas	0,44	0,45	-0,0141	0,6453	Não
estuda_noturno	0,56	0,55	0,0028	0,9264	Não
perc_rfp_ate_1salmeio_curso	70,53	70,15	0,0249	0,4198	Não
genero_masculino	0,40	0,40	-0,0067	0,8278	Não
cor_branca	0,36	0,35	0,0107	0,7261	Não
tempo_entre_conclusao_em_ingresso	4,42	4,24	0,0358	0,2391	Não

Fonte: Elaborada pelo autor (2024)

Os gráficos da Figura 63 conferem validade à afirmação anterior sobre o balanceamento dos dados, expondo o exemplo de três das sete variáveis independentes que compõem o modelo do PSM (*grau_academico_licenciatura*, *tempo_entre_conclusao_em_ingresso* e *perc_rfp_ate_1salmeio_curso*). Nota-se, nos gráficos gerados com o conjunto completo de dados, considerável dispersão nos casos, enquanto nos gráficos dos dados pareados os casos dispersos se concentram nas extremidades para a primeira variável e próximos à linha de normalidade na segunda e terceira variáveis.

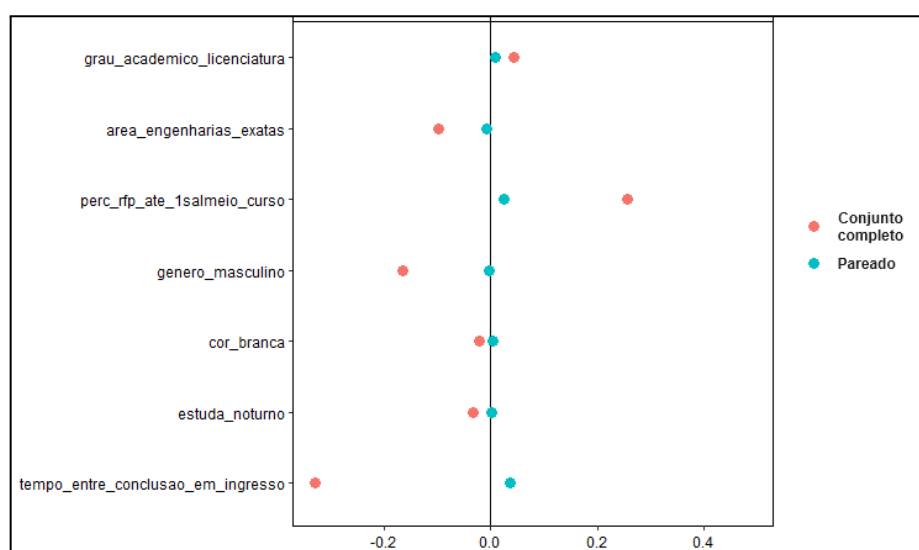
Figura 63 - Distribuição dos dados nos conjuntos completo e pareado



Fonte: Elaborada pelo autor (2024)

A Figura 64 apresenta, para cada variável independente, a diferença da média do escore de propensão entre os grupos de tratamento e de controle no conjunto com todos os dados (completo) e no pareado, confirmando que o processo de *matching* foi executado com coerência e proveu, como principal resultado, um conjunto de dados composto por estudantes bolsistas e não bolsistas portadores de características muito similares.

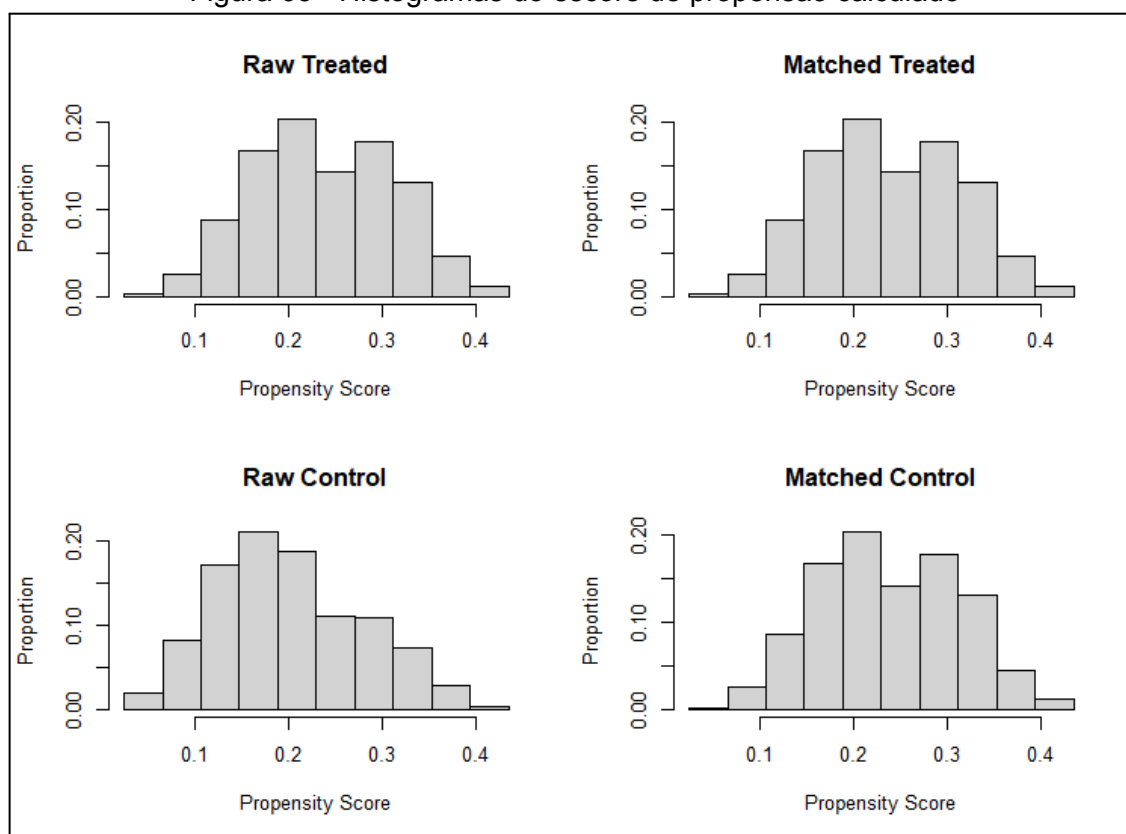
Figura 64 - Diferença do escore de propensão das variáveis independentes



Fonte: Elaborada pelo autor (2024)

Outra forma de analisar o balanceamento entre os grupos de controle e de tratamento é comparando a distribuição do escore de propensão calculado nos dois conjuntos de dados (completo e pareado). Observa-se, nos dois gráficos superiores da Figura 65, que a distribuição é a mesma entre os estudantes do grupo de tratamento nos dois conjuntos de dados (*raw* refere-se ao conjunto completo e *matched* ao pareado). Entretanto, ao analisar a distribuição nos gráficos inferiores, relativos ao grupo de controle, fica visível como eles seguem uma distribuição normal no conjunto pareado, além de refletir uma distribuição muito próxima ao grupo de tratamento.

Figura 65 - Histogramas do escore de propensão calculado



Fonte: Elaborada pelo autor (2024)

Uma vez que o conjunto de dados pareado está com o devido balanceamento, isto é, as variáveis independentes dos estudantes do grupo de controle possuem valores iguais ou próximos àquelas dos estudantes do grupo de tratamento, é possível estimar o impacto do recebimento do auxílio socioeconômico no desempenho acadêmico e na evasão. Optou-se por utilizar duas medidas para indicar este impacto, o Efeito Médio de Tratamento sobre o Tratado (ATT) e a diferença das médias entre os dois grupos.

Os resultados dispostos na Tabela 42 confirmam, no âmbito dos cursos superiores do IFMG, o positivo impacto do auxílio socioeconômico sobre o desempenho acadêmico e a evasão dos estudantes. No caso do desempenho acadêmico, na medida do ATT e da diferença das médias, o fato de receber o auxílio faz com que o coeficiente de rendimento seja incrementado, na média, em pouco mais de 17 pontos. Referente à evasão, o ATT mensura que o auxílio promove uma proporção menor de 21,10%, enquanto na diferença das médias essa proporção é ainda maior (22,74%).

Tabela 42 - Impacto do auxílio socioeconômico sobre desempenho e evasão

Construto	ATT	<i>p</i>	Diferença das médias	<i>p</i>
Desempenho acadêmico	17,58	< 0,0001 (significativo)	17,82	< 0,0001 (significativo)
Evasão	-21,10%	< 0,0001 (significativo)	-22,74%	< 0,0001 (significativo)

Fonte: Elaborada pelo autor (2024)

Tais resultados constataam a importância deste recurso para melhorar os indicadores institucionais ligados ao desempenho acadêmico e à evasão. Embora o atendimento do auxílio socioeconômico não chegue a 20% dos estudantes matriculados em cursos de graduação do IFMG, resta comprovada a contribuição direta da política pública do PNAES na otimização de dois aspectos fundamentais para as instituições educacionais, assim como previsto no próprio escopo do Programa.

5 CONSIDERAÇÕES FINAIS

As tecnologias digitais conseguiram, em um curto espaço de tempo quando comparadas com outras inovações modernas, revolucionar a forma como a sociedade interage com os dados eletrônicos, fazendo com que um volume cada vez maior de informações seja produzido e consumido sob os mais variados meios. As organizações públicas estão inseridas neste contexto a partir do momento em que sua forma de trabalhar passa a depender do acesso a informações de qualidade, com capacidade para contribuir com os processos decisórios e influenciar ações. Este é um dos papéis da área de CI, fazer com que a estrutura organizacional esteja adequada para controlar o fluxo informacional de maneira que os processos sejam executados com eficiência. Para isso, a CI interage com outras áreas de conhecimento no intuito de aplicar as teorias e práticas que colaborem com o alcance de seus objetivos.

A elaboração de um estudo acadêmico envolvendo diversas áreas de conhecimento pode não ser uma tarefa simples, mas ela se torna factível aliando um adequado escopo e aprofundamento das bases teóricas com uma metodologia coerente com os objetivos propostos. Nesta tese buscou-se construir um referencial teórico amplo, de forma que os vários assuntos relacionados com o tema central fossem abordados sob uma perspectiva caracterizadora, isto é, que sirva não somente para promover conceituações, mas que habilite o leitor a compreender o contexto do assunto e suas particularidades. A partir da elaboração deste referencial, foi possível desenvolver novos conhecimentos sobre os três construtos que originaram a demanda materializada na investigação, bem como acerca de outros temas necessários para a consecução das atividades propostas como percurso metodológico.

Dado esse contexto, a metodologia proposta envolveu a utilização de diferentes técnicas e ferramentas de acordo com as particularidades de cada objetivo a ser alcançado. Optou-se pelo emprego de recursos alinhados com o que se pretendia fazer com a forma mais adequada de execução. Outro aspecto relevante é que houve a descrição dos procedimentos metodológicos de maneira a promover a sua potencial reprodutibilidade em outros cenários, como no caso de uma IFES com objetivos semelhantes.

A proposta principal apresentada buscou investigar e analisar, sob uma ótica orientada a ontologias, quais são as relações entre a distribuição de auxílio socioeconômico de uma IFES, a evasão escolar e o desempenho acadêmico dos estudantes. Dada sua amplitude, em especial por envolver assuntos com particularidades relevantes que merecem ser tratados separadamente antes de integrados, foi necessário dividi-la para viabilizar a estruturação em objetivos específicos das atividades necessárias para o seu completo cumprimento. Assim sendo, a primeira seção deste Capítulo apresenta as considerações

finais relacionadas a cada um dos seis objetivos específicos que guiaram as atividades da tese e, ao terem seus resultados integrados, possibilitaram o cumprimento dos dois objetivos gerais propostos.

5.1 Considerações quanto aos objetivos específicos

Quanto ao primeiro objetivo específico concluído, apresentou-se um exemplo real sobre como um ambiente estruturado com base nos preceitos da Gestão da Informação contribui significativamente para a recuperação de dados que, mediante aplicação de técnicas para o seu processamento com um objetivo bem definido, podem se transformar em informações úteis aos gestores em processos de tomada de decisão.

No caso em questão, a caracterização do perfil das solicitações de auxílio socioeconômico à assistência estudantil do IFMG somente foi viável graças ao prévio trabalho institucional de organização do fluxo informacional relativo ao processo que trata da concessão de tais benefícios. Muito dessa estruturação passa pela disponibilização de um sistema informatizado para uso pelos estudantes e servidores da instituição vinculados a este processo, o qual possibilitou que os dados de todos os *campi* fossem mantidos de forma integrada no mesmo repositório, habilitando, por conseguinte, a sua recuperação e manuseio nesta tese. Ademais, trata-se de um estudo cujos resultados podem servir de base para discussões no nível institucional quanto à pertinência de efetivar alterações na PAE do IFMG, espelhando a realidade do público que busca a assistência estudantil como apoio para a continuidade dos seus estudos.

Outra tarefa no âmbito do primeiro objetivo consistiu na aplicação de uma estratégia de *clustering* visando identificar grupos com características similares dentre os participantes de processos seletivos da assistência estudantil. Esta estratégia permitiu explorar os dados brutos coletados com um formato diferente daquele que geralmente é utilizado em estudos com foco similar a este, nos quais os valores de todas as variáveis da amostra são apresentados quantitativamente para informar, e não comparar. Ao construir perfis com base nas características comuns de um mesmo grupo de elementos, a estratégia empregada conseguiu mostrar que em um único conjunto de dados existem três tipos diferentes de indivíduos que demandam apoio da assistência estudantil para a continuidade dos seus estudos, fato que deve ser analisado pelos gestores envolvidos com o processo de seleção dos bolsistas para que a distribuição dos recursos contemple a realidade do público da instituição.

O segundo objetivo concluído gerou como resultado um conjunto de informações que caracterizam o comportamento informacional dos servidores que atuam diretamente na concessão de auxílios socioeconômicos no IFMG. A exploração dos dados coletados mostrou que existia, naquele momento, certa preferência entre os participantes da pesquisa

em utilizar recursos de TI quando buscam e compartilham informações. Tal fato sugere que o próprio fluxo informacional da instituição privilegia a adoção de tais recursos em contraponto a um formato que favoreça o contato pessoal entre os indivíduos. Além disso, verificou-se que os tipos de informações mais usados são aderentes com o perfil profissional dos servidores em questão, assim como também possuem relação direta com as principais necessidades informacionais dos mesmos.

Com os resultados desta pesquisa, a instituição pode compreender melhor como é o trato dos servidores do setor de assistência estudantil com a informação e, a partir de tal entendimento, propor aperfeiçoamentos que se traduzam em formas mais otimizadas de executar as ações necessárias para o bom andamento do processo de análise socioeconômica dos demandantes de auxílios.

Conforme previsto, o cumprimento desses dois objetivos proporcionou conhecimentos ampliados sobre a realidade da assistência estudantil do IFMG, tanto em relação aos estudantes e servidores envolvidos, quanto acerca dos dados mantidos no sistema informatizado de gestão dos processos seletivos. Este maior entendimento contribuiu diretamente na etapa seguinte, a elaboração da ontologia, fazendo com que o artefato construído estivesse adequado à realidade institucional.

O foco do terceiro objetivo específico era executar atividades que, ao final, produzissem como resultado uma ontologia de domínio para organizar os conceitos e integrar os dados dos três construtos analisados no âmbito do IFMG. Enquanto suporte metodológico, a abordagem SABiO colaborou para a estruturação do fluxo de trabalho com a indicação de processos que, por sua vez, simplificaram o esforço de produção da ontologia sob o ponto de vista procedimental. É importante ressaltar que a elaboração da ontologia contou com a participação de especialistas externas com diferentes graus de conhecimento sobre os construtos envolvidos nesta pesquisa, fato que indica a pacificação dos conceitos e elementos constantes na versão final publicada da ontologia.

No que tange ao artefato construído, ele está alinhado com o perfil da OntoUML e, por conseguinte, atende aos pressupostos da ontologia de fundamentação UFO. Isso permitiu validar os aspectos metafísicos que caracterizam a natureza profunda dos seres representados pela ontologia. Pesquisadores da Universidade de Twente aprovaram a inclusão da Onto-Educacional no catálogo¹² colaborativo de modelos OntoUML/UFO. Esse catálogo serve como principal referência para a pesquisa empírica e de modelagem conceitual em OntoUML e UFO.

No âmbito do quarto objetivo específico, buscou-se construir um conjunto de dados composto por instâncias reais do IFMG que representassem os elementos da ontologia. Merece destaque a superação dos obstáculos para a integração de dados provenientes de

¹² <https://github.com/OntoUML/ontouml-models/tree/master/models/melo2024onto-educational>

diferentes fontes e sistemas, visando à padronização e uniformização conforme a ontologia operacional estabelecida. A construção de um repositório intermediário mostrou-se uma escolha acertada, dada a possibilidade de execução de diversas rotinas de pré-processamento dos dados em um local independente e adequado à estrutura da ontologia, bem como ao enriquecimento promovido pela adição de dados advindos de repositórios externos relacionados ao contexto da pesquisa. Por fim, o mapeamento entre os dados e os conceitos da ontologia possibilitou a transformação desses dados para o formato exigido pela plataforma de *Graph Database*, responsável por prover um ambiente para obtenção das respostas às questões de competência e, assim, contribuir para avaliar a ontologia sob o formato de verificação.

O *dashboard* gerencial, elaborado como cumprimento do quinto objetivo específico, não somente proporcionou uma visualização mais atraente e integrada dos dados armazenados na estrutura de grafos, mas principalmente subsidiou a realização de análises sobre as relações entre os construtos pelo pesquisador e pelas especialistas de domínio.

No contexto do último objetivo específico, há de se destacar a aplicação de uma técnica, validada por outras publicações científicas, capaz de identificar relações de causalidade entre políticas públicas e a sociedade, tal como se pretendia ao investigar o impacto do PNAES no IFMG. Embora todas as fases do ciclo de políticas públicas possuam particularidades que, por vezes, tornam o processo mais complexo, as atividades de avaliação merecem destaque pela sua capacidade de orientar a condução dos trabalhos, independentemente do estágio em que eles se encontram. As análises promovidas com base nos resultados das avaliações são fundamentais para que não ocorram desvios de finalidade da política, agregando novas informações que, por sua vez, prestam contas à sociedade de como os recursos públicos vêm sendo consumidos. Os resultados obtidos indicaram a efetividade do PNAES na instituição, seja enquanto auxílio para a redução da evasão dos estudantes atendidos ou como fomentador de uma realidade com melhores condições para que os bolsistas obtenham desempenho superior em relação àqueles não atendidos.

A Seção seguinte introduz algumas considerações do pesquisador sob os pontos de vista socioeconômico e técnico, respondendo a questão que norteou a elaboração da tese e indicando o cumprimento dos objetivos gerais da pesquisa.

5.2 Considerações quanto aos objetivos gerais

Cumpridos os objetivos específicos, têm-se o consequente alcance dos dois objetivos gerais definidos na tese. Os resultados das atividades conduzidas para o alcance do primeiro deles, voltado para uma análise denominada “socioeconômica”, foram amplamente explorados ao longo do Capítulo 4, no qual buscou-se apresentar indicativos da

existência de vínculos diretos entre a distribuição de auxílio socioeconômico, a evasão escolar e o desempenho acadêmico dos estudantes do IFMG.

De maneira geral, é necessário que as IFES fomentem a realização de processos decisórios baseados em evidências, seja no âmbito administrativo ou acadêmico, ao invés de tomar ações amparadas somente por percepções subjetivas. Estudos como o desta tese visam contribuir nesse sentido, demonstrando quantitativamente que o investimento em ações de atendimento a estudantes com vulnerabilidades socioeconômicas gera efeitos positivos em dois importantes indicadores de gestão educacional, o desempenho acadêmico dos estudantes e a evasão escolar. No caso dessa última, fortemente impactada nas instituições após a pandemia de COVID-19, a tese se mostra ainda mais alinhada, podendo servir como suporte para a construção de novas políticas de atendimento aos estudantes. Ademais, tem-se um avanço nos estudos sobre o impacto causado pelo PNAES nas IFES, dada a ausência de uma regulação formal para a avaliação do alcance dos objetivos do programa.

Considerando que, dentre os dados do IFMG analisados com o devido rigor estatístico, a média do desempenho acadêmico dos bolsistas é 32% superior à média dos não bolsistas, a proporção de evadidos é 43,5% menor entre os bolsistas e a proporção de formados é duas vezes superior também entre aqueles que recebem auxílio institucional, a conclusão do primeiro objetivo geral (com enfoque socioeconômico) é que, sim, a distribuição do auxílio está diretamente relacionada com a redução da evasão e aumento do desempenho acadêmico dos estudantes de cursos de graduação. Embora o percentual de bolsistas seja pequeno em comparação ao universo total de estudantes do IFMG, situação ocasionada pela insuficiência de recursos financeiros, resta comprovado que dois objetivos do PNAES são alcançados na instituição, além de outros que são beneficiados positivamente, tais como o aumento nas taxas de retenção e a redução das desigualdades sociais regionais.

O segundo objetivo geral tem um aspecto mais “técnico” ao propor a análise, sob o ponto de vista subjetivo do pesquisador, de eventuais contribuições da abordagem ontológica no escopo das tarefas de integração e processamento de dados obtidos em fontes heterogêneas. A alternativa metodológica escolhida para a realização desta tarefa foi a fenomenologia, caracterizada por Goldenberg (2011) como um método, no sentido de atitude, que requer uma postura própria do pesquisador para que seus conhecimentos prévios, emoções, teorias e suposições não interfiram na sua compreensão acerca da essência do fenômeno investigado. O que foi compreendido advém da observação promovida durante as experiências do pesquisador envolvendo o fenômeno em questão, de maneira que a análise não está focada na aparência imediata do fenômeno, mas na sua verdade.

O primeiro aspecto a ser considerado como uma contribuição da ontologia corresponde à formalização dos conceitos do domínio trabalhado, fato que promoveu consistência ao estudo na medida em que foi possível garantir que o significado de cada conceito se manteve íntegro ao longo da execução das atividades que envolveram o uso dos dados coletados após a elaboração da Onto-Educacional. A interpretação dos dados também se beneficiou deste aspecto, pois as afirmações indicadas ao longo das análises se referiram a conceitos que, sob o ponto de vista do pesquisador e das especialistas de domínio do IFMG, estavam livres de ambiguidades na sua definição. É possível conceber, portanto, que a ontologia contribuiu para representar formalmente o conhecimento existente em um contexto delimitado das áreas de evasão escolar, desempenho acadêmico e distribuição de auxílio socioeconômico.

A representação gráfica da modelagem ontológica traz maior clareza no sentido da análise visual dos elementos que a compõem, em especial quando comparada com a representação proposta pela modelagem clássica, por exemplo, a modelagem física de um banco de dados relacional. Isto se traduz em uma interessante possibilidade para atuação em conjunto com os especialistas de domínio que fazem parte do projeto, pois a representação do modelo em um formato mais acessível permite que o próprio artefato sirva como base para discussões sobre ajustes e melhorias necessárias para adequá-lo à realidade que está sendo representada pela ontologia. Destaca-se, ainda, a flexibilidade do modelo ontológico para representar conexões complexas sem a necessidade de promover alterações significativas na sua estrutura ou lançar mão do emprego de múltiplas junções entre os elementos.

Ainda sob esse aspecto da compreensão simplificada de elementos gráficos, é importante ressaltar a visão intuitiva promovida pelos grafos em contrapartida da visão tabular obtida em sistemas baseados no modelo relacional. As relações entre os elementos que compõem a ontologia podem ser visualizadas sob variados prismas na estrutura baseada em grafos, o que na prática se transforma em informações com maior valor agregado aos usuários do que aquelas expostas em linhas e colunas. Novos conhecimentos podem ser mais facilmente construídos por meio da compreensão das relações apresentadas em um grafo de forma mais amigável.

A flexibilidade do modelo ontológico se evidenciou quando surgiu a demanda pela integração dos dados particulares da instituição (acadêmicos e da assistência estudantil) com dados externos provenientes de outros repositórios, tendo como intuito o enriquecimento do conjunto de dados. Esse processo exigiu pequenas adaptações e expansões do modelo inicialmente elaborado, permitindo que diferentes fontes de dados, com estruturas e vocabulários distintos, fossem harmonizadas de forma coerente. A capacidade do modelo ontológico de incorporar novas classes, propriedades e relações, se

mostrou essencial para garantir a interoperabilidade entre sistemas heterogêneos e atender às necessidades de integração de dados para a pesquisa conduzida nesta tese.

Os benefícios de uso da ontologia também se fizeram notar quando do armazenamento dos dados no repositório baseado em grafos. Sob o ponto de vista prático, a facilidade de entendimento proporcionada pelo mapeamento entre o conjunto de dados disponível no repositório intermediário e os elementos da ontologia, fizeram com que os comandos escritos para a ingestão dos dados no repositório em grafo tivessem sintaxe simplificada e, conseqüentemente, demandassem curto tempo para execução. Ademais, é válido ressaltar que o formato para representar o uso dos dados na abordagem ontológica é mais próximo ao entendimento humano do que o formato de inúmeras junções requerido por modelos relacionais tradicionais.

No que se refere à recuperação das informações, a ontologia colaborou para que as consultas escritas fossem simplificadas, objetivas e de fácil entendimento até mesmo por pessoas sem conhecimento técnico avançado da linguagem utilizada. Ao contrário de consultas escritas em sistemas baseados no modelo relacional, as consultas elaboradas para responder as questões de competência desta tese não demandaram a especificação de sucessivas junções entre cadeias de elementos para obter as informações desejadas. Percebeu-se, ainda, que o próprio modelo ontológico tornou mais intuitivo o processo de escrita das consultas, pois a distribuição visual dos elementos no seu escopo, por si só, direcionou o entendimento acerca de como deveria ser a estruturação de cada consulta.

Como último aspecto de contribuição da ontologia elaborada nesta tese, evidencia-se o fato dela ter suportado a construção de um mesmo conjunto de dados apto a diversificadas utilizações, como a construção do repositório em grafos, a elaboração do *dashboard* gerencial com recursos de visualização de informações para análise das relações entre os construtos da pesquisa, além da aplicação do método *Propensity Score Matching* para análise do impacto do PNAES no IFMG. Essa característica de compartilhamento do significado dos conceitos representados pelos dados, não somente das instâncias em si, confere alta recomendação ao uso de ontologias em projetos que necessitem de colaboração e interoperabilidade entre sistemas e diferentes indivíduos.

Respondendo a questão de pesquisa que norteou o desenvolvimento da tese, apoiado nas constatações subjetivas do pesquisador, a abordagem ontológica favoreceu positivamente a construção de uma estrutura informacional para integrar os dados dos três construtos, possibilitando que as relações entre eles fossem investigadas sob diferentes perspectivas. Trata-se, também, de uma abordagem que incrementa a qualidade do processamento de dados em projetos que requerem a representação formal de um domínio do conhecimento, pois proporciona uma base sólida com um modelo semântico claro e

consensuado, próprio para a aplicação de técnicas avançadas de variadas áreas como, por exemplo, a Ciência de Dados.

Embora a tese abarque várias áreas de conhecimento, trata-se de um estudo conduzido no âmbito da Ciência da Informação, motivo pelo qual na próxima Seção são mencionadas as contribuições da tese para a CI de forma geral.

5.3 Contribuições para a área de Ciência da Informação

Pontuar eventuais aspectos contribuidores de uma tese para área tão ampla como a Ciência da Informação é, no mínimo, uma tarefa de complexa execução, porém plausível, dada sua característica de interligação com outras áreas de conhecimento. É preciso reconhecer, inicialmente, que os artefatos gerados ao longo da consecução desta tese são, em sua maioria, produtos informacionais que promoveram novos conhecimentos ou poderão, a partir de sua utilização por terceiros, colaborar com o entendimento de outras realidades.

O estudo realizado sobre o comportamento informacional dos servidores que atuam no processo de concessão de auxílios socioeconômicos avança nesta importante subárea da CI que visa compreender melhor como os indivíduos se relacionam com a informação sob determinadas condições e contextos. No caso em questão, escassos são os estudos que analisam o comportamento informacional para o público citado, de maneira que os resultados obtidos na tese passam a servir de referência em futuras pesquisas similares. Outro aspecto inovador refere-se à elaboração do estudo sobre o comportamento informacional antes do início da construção da ontologia, fato que permitiu, sob o ponto de vista do pesquisador, compreender previamente as reais necessidades de informação de um público com interesse direto nos resultados da pesquisa, alinhando os conceitos do artefato ontológico com a realidade prática vivida por profissionais da área de assistência estudantil.

O caráter transdisciplinar da CI é reforçado pelos resultados desta tese, dado que em seu percurso diversas interações com outras áreas de conhecimento ocorreram para que os objetivos propostos fossem alcançados. Raramente os produtos informacionais construídos no espectro de um projeto de CI existem por si só, abarcados exclusivamente no âmbito dessa ciência, mas são projetados e desenvolvidos com o apoio de métodos e técnicas compartilhadas por áreas que se complementam. Como exemplo, é possível mencionar a aplicação de técnicas advindas da Estatística, seja para promover análises caracterizadoras de dados brutos ou para investigar, confirmando ou não, a existência de relações entre os construtos da pesquisa. A área de Ciência de Dados também forneceu importantes recursos para as atividades da tese, dentre as quais se evidenciam as técnicas de aprendizado de máquina, visualização de dados e integração de fontes de dados heterogêneas. A área de Políticas Públicas, envolvida nesta tese pela investigação sobre o impacto do PNAES na

instituição pesquisada, também se relaciona com a área de CI, em especial no que tange a questões de avaliação dos resultados gerados pelas ações da política. A tese avança neste sentido ao mostrar que um ambiente organizado informacionalmente nos moldes balizados pela CI é capaz de prover os recursos necessários para a aplicação de técnicas que vão corroborar os resultados da política.

Outra contribuição da tese para a área de CI consiste na apresentação prática de como foi realizado todo o ciclo de vida de construção de uma ontologia de domínio. Como mencionado anteriormente, o processo de desenvolvimento de ontologias não é trivial, seja pelas características particulares do contexto ou da necessidade de integração com outras áreas de conhecimento, em especial a Computação. Sendo assim, o trabalho exposto na tese permite compreender, em detalhes, como ocorreu a execução de cada etapa da construção do artefato ontológico, o planejamento de questões relacionadas com o armazenamento dos dados, o emprego da ontologia operacional como base para o processamento de máquina, a técnica para recuperação de informações e sua apresentação em um formato atrativo para os usuários. Tem-se, portanto, um exemplo que pode servir de base prática para estudos similares, reduzindo a curva de aprendizado da criação e uso de artefatos ontológicos como suporte para integração semântica de dados de fontes heterogêneas.

Por fim, é válido mencionar a interligação da tese ao paradigma físico da CI, especialmente por lidar com a organização, a estruturação e a representação do conhecimento. No contexto físico, a ontologia é usada para estruturar e modelar dados e informações de maneira que as relações e as propriedades de entidades no mundo real (objetos físicos, fenômenos naturais, processos) sejam claramente definidas e compreendidas. A ontologia elaborada nesta tese:

- favorece o entendimento sobre a definição de conceitos e relações, além de sua organização hierárquica e taxonômica (representação formal do conhecimento);
- ilustra a importância da precisão na interpretação dos dados ao definir o significado formal (isto é, a semântica formal) das entidades e das suas interações;
- contribui para o emprego de recursos inteligentes no processo de busca das informações (sistemas de recuperação de informação);
- facilita a interoperabilidade de dados e sistemas ao definir os padrões de informações (interoperabilidade e padrões).

A próxima Seção discute algumas dificuldades enfrentadas ao longo do desenvolvimento da tese e propõe sugestões de trabalhos complementares ao exposto.

5.4 Dificuldades encontradas e trabalhos futuros

O fator inicial que ocasionou dificuldades para a realização deste trabalho está relacionado com a amplitude dos três construtos abordados, bem como das interações existentes entre eles. Cada construto possui suas próprias conceituações e contextos de aplicação, de maneira que a delimitação do escopo do trabalho precisou ser avaliada em diversos momentos do trabalho para que as atividades propostas para alcance dos objetivos não ficassem demasiadamente complexas.

A compreensão do fluxo mais adequado para elaboração da ontologia também trouxe desafios durante o desenvolvimento da tese, pois não se trata de um processo trivial, ainda mais com a demanda de utilização da ontologia como base para a integração semântica dos dados, e não somente como referencial conceitual. Neste sentido, a metodologia SABiO contribuiu diretamente para o entendimento das etapas necessárias ao processo de construção da ontologia, servindo de referência para a construção do conjunto de atividades executadas no âmbito dos procedimentos metodológicos da tese.

Ainda sobre a elaboração da ontologia, é importante pontuar que não é simplificado o entendimento dos preceitos da UFO para adequação da ontologia de referência, o que demandou esforço adicional de estudos para que o artefato construído estivesse coerente com o proposto na ontologia de fundamentação.

A integração de recursos da área de Ciência de Dados ao processo de criação da ontologia demandou atenção especial para que os resultados de sua aplicação fornecessem subsídios para a elaboração dos artefatos ontológicos. As técnicas aplicadas foram selecionadas em função daquilo que poderiam contribuir no sentido de retratar, por meio dos dados coletados, a realidade da instituição pesquisada.

Outra dificuldade na pesquisa foi a inexistência de uma base legal que estabeleça critérios objetivos para a avaliação da efetividade do PNAES enquanto política pública aplicada nas IFES. A ausência de indicadores padronizados para mensurar o impacto das ações do programa na redução da evasão e no desempenho acadêmico dos estudantes em situação de vulnerabilidade socioeconômica compromete a comparabilidade dos resultados e a identificação de boas práticas que, por sua vez, poderiam ser replicadas em diferentes contextos. Os indicadores quantitativos utilizados nesta tese permitiram estimar o impacto da política no contexto do IFMG e servem como evidências para o aprimoramento das ações de assistência estudantil na instituição.

Diante das limitações e alcances desta pesquisa, algumas sugestões para trabalhos futuros podem contribuir com o aprofundamento da análise sobre as relações entre os três construtos. Embora o foco nesta tese tenha sido estudantes de cursos de graduação, o modelo metodológico desenvolvido está habilitado para aplicação em outros níveis de ensino, como os cursos técnicos. Considerando que a realidade socioeconômica dos

estudantes desses cursos muitas vezes se assemelha à dos graduandos, investigações futuras poderiam explorar se os padrões identificados também se manifestam nesse público. Tal abordagem permitiria não somente uma ampliação do escopo de análise de efetividade dos auxílios estudantis, mas também forneceria subsídios para a formulação de políticas públicas mais abrangentes e direcionadas às especificidades dos diferentes níveis de ensino.

Outra possibilidade para trabalhos futuros está na superação da limitação imposta pela dependência de variáveis previamente disponibilizadas pelo IFMG na construção do conjunto de dados da ontologia. Como a presente pesquisa foi condicionada às informações existentes nos registros institucionais, não foi possível expandir o conjunto de características pessoais dos estudantes analisados. Estudos futuros poderiam buscar a incorporação de novas variáveis, obtidas por meio de fontes complementares, como questionários aplicados diretamente aos alunos, dados socioeconômicos mais detalhados ou informações qualitativas sobre fatores subjetivos que influenciam a evasão e o desempenho acadêmico. A ampliação desse escopo de dados permitiria uma modelagem mais abrangente e uma compreensão mais profunda das dinâmicas envolvidas na permanência estudantil, possibilitando a formulação de estratégias de assistência mais eficazes e personalizadas.

Por fim, outra sugestão de trabalhos futuros é aprofundar a investigação sobre a contribuição do estudo do comportamento informacional para o desenvolvimento da ontologia, avaliando se essa etapa deveria ser incorporada como parte essencial do processo de modelagem ontológica. Pesquisas adicionais podem explorar metodologias capazes de integrar a análise do comportamento informacional na fase inicial da construção de ontologias, verificando se essa abordagem contribui para uma modelagem mais precisa e alinhada às necessidades dos usuários da informação.

5.5 Publicações dos resultados obtidos

O desenvolvimento das atividades desta tese possibilitou a elaboração de sete artigos científicos para periódicos científicos, dos quais cinco foram publicados e dois estão em processo de avaliação, além de um artigo completo publicado nos anais do principal evento sobre ontologias no Brasil, a ONTOBRAS. Tal resultado confirma, sob o ponto de vista externo, o impacto positivo para a geração e difusão de conhecimentos proporcionado pela pesquisa realizada, além da contribuição para o Programa ao qual ela está vinculada (Programa de Pós-Graduação em Gestão & Organização do Conhecimento da Universidade Federal de Minas Gerais).

Artigos completos publicados em periódicos:

- [Visualização dos dados de candidatos a auxílios socioeconômicos em grafos: estudo de caso no Instituto Federal de Minas Gerais \(IFMG\)](#)
Autores: Eduardo Cardoso Melo, Marcello Peixoto Bax e Yuri Bento Marques
Periódico: Revista AtoZ - Novas Práticas em Informação e Conhecimento
Ano: 2022
- [Perfil de comportamento informacional dos servidores envolvidos com o processo de concessão de auxílios socioeconômicos em uma Instituição Federal de Ensino Superior](#)
Autores: Eduardo Cardoso Melo e Marcello Peixoto Bax
Periódico: Ciência da Informação em Revista
Ano: 2024
- [Gestão da Informação como suporte para a caracterização das solicitações de auxílios socioeconômicos na Assistência Estudantil do Instituto Federal de Minas Gerais \(IFMG\)](#)
Autores: Eduardo Cardoso Melo, Marcello Peixoto Bax, Paulo César Lourenço da Silva, Yuri Bento Marques e Evaldo de Oliveira da Silva
Periódico: Revista AtoZ - Novas Práticas em Informação e Conhecimento
Ano: 2024
- [Identificação de perfis das solicitações de auxílios socioeconômicos com a técnica de clustering: estudo de caso no Instituto Federal de Minas Gerais \(IFMG\)](#)
Autores: Eduardo Cardoso Melo e Marcello Peixoto Bax
Periódico: Revista InCid
Ano: 2024
- [Caracterização da assistência estudantil no IFMG e o seu papel na construção de possibilidades para os estudantes](#)
Autores: Eduardo Cardoso Melo e Marcello Peixoto Bax
Periódico: Educação On-line PUC-Rio
Ano: 2024
- [Uma ontologia para integrar dados de evasão escolar, desempenho acadêmico e distribuição de auxílios socioeconômicos: estudo de caso no IFMG](#)
Autores: Eduardo Cardoso Melo e Marcello Peixoto Bax

Periódico: Perspectivas em Ciência da Informação

Ano: 2025

Artigos completos em processo de avaliação:

- Mensuração do impacto do PNAES no desempenho acadêmico e na evasão escolar do IFMG

Autores: Eduardo Cardoso Melo e Marcello Peixoto Bax

Periódico: Administração Pública & Gestão Social

Artigos completos publicados em anais de eventos:

- [Uma ontologia para integrar dados de evasão escolar, desempenho acadêmico e distribuição de auxílios socioeconômicos: estudo de caso no IFMG](#)

Autores: Eduardo Cardoso Melo e Marcello Peixoto Bax

Evento: Seminário de Pesquisa em Ontologias no Brasil (ONTOBRAS)

Ano: 2024

REFERÊNCIAS

- ABREU, M. K. A.; XIMENES, V. M. Pobreza, permanência de universitários e assistência estudantil: uma análise psicossocial. **Psicologia USP**, v. 32, 2021. Disponível em: <https://www.scielo.br/j/pusp/a/pBtyBfxJqkXbvzwVvcQprzS/?lang=pt>. Acesso em: 9 jul. 2022.
- AGGARWAL, C. C. An introduction to cluster analysis. **Data Clustering**. Chapman and Hall/CRC, 2018. p. 1-28. Disponível em: <https://www.taylorfrancis.com/chapters/edit/10.1201/9781315373515-1/introduction-cluster-analysis-charu-aggarwal>. Acesso em: 31 ago. 2022.
- ALMEIDA, F. M. S.; CASSUCE, F. C. C.; CIRINO, J. F. Desempenho acadêmico na Universidade Federal de Viçosa. **Revista de Ciências Humanas**, v. 16, n. 1, p. 155-175, 2016. Disponível em: <https://www.locus.ufv.br/handle/123456789/12970>. Acesso em: 03 ago. 2022.
- ALMEIDA, G. S.; LOPES, J. M. R. O que explica o desempenho dos estudantes? Um estudo de revisão. **Revista de Instrumentos, Modelos e Políticas em Avaliação Educacional**, v. 1, n. 1, p. e020004-e020004, 2020. Disponível em: <https://revistas.uece.br/index.php/impa/article/view/3597>. Acesso em: 20 jul. 2022.
- ALMEIDA, J. P. A. et al. Conceptual Modeling. 42nd International Conference on Conceptual Modeling, 2023. Disponível em: <https://research.utwente.nl/en/publications/conceptual-modeling-42nd-international-conference-er-2023-lisbon->. Acesso em: 24 abr. 2025.
- ALMEIDA, J. P. A. et al. **gUFO**: A Lightweight Implementation of the Unified Foundational Ontology (UFO). 2019. Disponível em: <http://purl.org/nemo/doc/gufo>. Acesso em: 06 nov. 2023.
- ALMEIDA, M. B.; BAX, M. P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Ciência da informação**, v. 32, p. 7-20, 2003. Disponível em: <https://www.scielo.br/j/ci/a/LR68syZsPSSmwwPHrNXmC8N/abstract/?lang=pt>. Acesso em: 20 set. 2022.
- ALMEIDA, M. B.; RIBEIRO, E. F.; MACULAN, B. C. Ontologies and Classification: the unavoidable interplay between human reasoning and machine reasoning. **Qualitative and Quantitative Methods in Libraries International Conference**, p. 1-13, 2019. Disponível em: https://mba.eci.ufmg.br/wp-content/uploads/Ontologies-and-Classification_subm.pdf. Acesso em: 22 mar. 2023.
- ALVES, J. M. A assistência estudantil no âmbito da política de educação superior. **Serviço Social em Revista**, v. 5, n. 1, 2002. Disponível em: http://www.uel.br/revistas/ssrevista/c_v5n1_Jo.htm. Acesso em: 2 jul. 2022.
- ALVES, K. S. **Evasão universitária**: consequências na vida pessoal do aluno. Trabalho de Conclusão de Curso. Universidade do Sul de Santa Catarina. Curso de Psicologia. 2008. Disponível em: <http://inf.unisul.br/~psicologia/wp-content/uploads/2008/07/Karolyne.pdf>. Acesso em: 02 mai. 2022.
- AMARAL, A. C. F. et al. Reflexos da COVID-19 sobre a evasão universitária na graduação: um estudo de caso de uma universidade pública da Região Centro-Oeste do Estado de Minas Gerais. **Humanidades e Tecnologia (FINOM)**, v. 34, n. 1, p. 83-96, 2022. Disponível

em: http://revistas.icesp.br/index.php/FINOM_Humanidade_Tecnologia/article/view/2207. Acesso em: 01 mai. 2022.

AMARAL, F. **Introdução à ciência de dados: mineração de dados e big data**. São Paulo: Alta Books, 2016.

ANDIFES. **V Pesquisa do Perfil Socioeconômico e Cultural dos Estudantes de Graduação das Instituições Federais de Ensino Superior Brasileiras**. Associação Nacional dos Dirigentes das Instituições Federais de Ensino Superior (ANDIFES). Brasília, 2019. Disponível em: <https://www.andifes.org.br/wp-content/uploads/2019/05/V-Pesquisa-do-Perfil-Socioecon%C3%B4mico-dos-Estudantes-de-Gradua%C3%A7%C3%A3o-das-Universidades-Federais-1.pdf>. Acesso em: 15 mar. 2022.

ANGELIM, I. C. S. Política de assistência estudantil na universidade pública: desafios para construção de uma concepção de direito de cidadania. **Seminário Políticas Sociais e Cidadania**, v. 3, 2010. Disponível em: http://www.interativadesignba.com.br/III_SPSC/arquivos/sessao7/199.pdf. Acesso em: 10 jun. 2022.

ANNEGUES, A. C.; PORTO JÚNIOR, S.; FIGUEIREDO, E. Tamanho da Turma e Desempenho Acadêmico dos Universitários: evidência para a UFPB. **Estudos Econômicos (São Paulo)**, v. 50, p. 99-124, 2020. Disponível em: <https://www.scielo.br/j/ee/a/tGSb8Tc9V57gnNWjS87WcmP/abstract/?lang=pt>. Acesso em: 23 ago. 2022.

ANTONIOU, G. et al. **A semantic web primer**. Third Edition. MIT press, 2012.

ARANHA, A. V. S.; PENA, C. S.; RIBEIRO, S. H. R. Programas de inclusão na UFMG: o efeito do bônus e do Reuni nos quatro primeiros anos de vigência-um estudo sobre acesso e permanência. **Educação em Revista**, v. 28, p. 317-345, 2012. Disponível em: <https://www.scielo.br/j/edur/a/BRCSC8y6dhMpQLwjVXgyskw/abstract/?lang=pt>. Acesso em: 10 jul. 2022.

ARAÚJO, A. C.; ANDRIOLA, W. B.; COELHO, A. A. Programa Institucional de Bolsa de Iniciação à Docência (PIBID): desempenho de bolsistas versus não bolsistas. **Educação em Revista**, v. 34, 2018. Disponível em: <https://www.scielo.br/j/edur/a/BKGWSsY3C6YPMrvzWmC7W3q/?lang=pt&format=html>. Acesso em: 04 ago. 2022.

ARAÚJO, E. P. O.; PAULA, C. P. A. Comportamento informacional: introdução de perspectivas simbólicas e afetivas em investigações sobre usuários de informação. **Prisma**, v. 34, p. 46-63, 2017. Disponível em: <http://aleph.letras.up.pt/index.php/prisma.com/article/view/3176>. Acesso em: 26 out. 2022.

ARAÚJO, M. V. M. **Sucesso e fracasso em projetos de tecnologia da informação: uma visão à luz dos fatores promotores e inibidores do alinhamento estratégico**. Dissertação de Mestrado. Universidade Federal de Pernambuco. Mestrado em Administração. 2016. Disponível em: <https://repositorio.ufpe.br/handle/123456789/25058>. Acesso em: 25 out. 2022.

ARAÚJO, R. A. E. F.; FERREIRA, M. A.; SILVA, V. R. Desempenho acadêmico e as variáveis socioeconômicas: um estudo com estudantes do curso de Ciências Contábeis. **Revista Científica Doctum Multidisciplinar**, v. 1, n. 5, 2022. Disponível em: <http://revista.doctum.edu.br/index.php/multi/article/view/378>. Acesso em: 30 jul. 2022.

ARAÚJO FILHO, A. A. **A sociedade do conhecimento e suas tecnologias**. São Paulo: Dialética, 2022.

ARGENTON, W. B. **Uma Arquitetura para Integração de Dados Educacionais no âmbito do ADE Norte Gaúcho**. Trabalho de Conclusão de Curso. Faculdade Meridional. Graduação em Ciência da Computação. 2020. Disponível em: <https://www.imed.edu.br/sobre-a-biblioteca-1/banco-de-tcc-s/ciencia-da-computacao-3>. Acesso em: 04 mar. 2022.

ARRETCHE, M. T. S. **Tendências no estudo sobre avaliação**. In: RICO, E. M. (org). Avaliação de Políticas Sociais: uma questão em debate. São Paulo: Cortez, 1998.

ASSIS, J. P.; SOUSA, R. P.; LINHARES, P. C. F. **Teste de Hipóteses Estatísticas**. Mossoró: EdUFERSA, 2020.

ASSUNÇÃO, R. C. **Antecedentes da evasão e permanência de alunos do ensino médio, técnico e superior**. Dissertação de Mestrado. Universidade FUMEC. Mestrado em Administração. 2021. Disponível em: <https://repositorio.fumec.br/handle/123456789/897>. Acesso em: 06 jul. 2022.

AUCELI, P.; BERARDI, R.; KOZIEVITCH, N. Integração semântica entre dados dos domínios da educação e segurança: um caso em Curitiba. **XV Escola Regional de Banco de Dados (ERBD)**, p. 141-150, 2019. Disponível em: <https://sol.sbc.org.br/index.php/erbd/article/view/8487>. Acesso em: 22 mar. 2023.

AVAZPOUR, I.; GRUNDY, J.; ZHU, L. Engineering complex data integration, harmonization and visualization systems. **Journal of Industrial Information Integration**, v. 16, p. 100103, 2019. Disponível em: <https://doi.org/10.1016/j.jii.2019.08.001>. Acesso em: 21 ago. 2022.

AYODELE, T. O. Types of machine learning algorithms. **New advances in machine learning**, v. 3, p. 19-48, 2010. Disponível em: https://www.researchgate.net/profile/Taiwo-Ayodele/publication/221907660_Types_of_Machine_Learning_Algorithms/links/53e53cdf0cf2fb7487156dac/Types-of-Machine-Learning-Algorithms.pdf. Acesso em: 29 ago. 2022.

BACK, L. B. Assistência estudantil: da formalização aos desafios na consolidação do direito. **Educação Online**, v. 14, n. 30, p. 32-52, 2019. Disponível em: <http://educacaoonline.edu.puc-rio.br/index.php/eduonline/article/view/510>. Acesso em: 5 jul. 2022.

BAGGI, C. A. S.; LOPES, D. A. Evasão e avaliação institucional no ensino superior: uma discussão bibliográfica. **Revista da Avaliação da Educação Superior (Campinas)**, v. 16, p. 355-374, 2011. Disponível em: <https://www.scielo.br/j/aval/a/RRGrQckrsd9CRGgKy4zkHXq/?lang=pt>. Acesso em: 23 set. 2022.

BAKER, J. Evaluating the impact of development projects on poverty: a handbook for practitioners. **Washington: The World Bank**, 2000. Disponível em: <https://documents.worldbank.org/pt/publication/documents-reports/documentdetail/470591468273625276/evaluating-the-impact-of-development-projects-on-poverty-a-handbook-for-practitioners>. Acesso em: 10 mai. 2024.

BARBOSA, A. T.; FERREIRA, E. B.; REZENDE, M. L. Indicadores para avaliação do Programa Nacional De Assistência Estudantil (PNAES). **Economia e Desenvolvimento**, v.

32, p. e8-e8, 2020. Disponível em: <https://periodicos.ufsm.br/eed/article/download/39740/pdf>. Acesso em: 10 jul. 2022.

BARBOSA, D. M. **Um modelo de gestão da informação e do conhecimento para o contexto da avaliação de cursos de graduação**. Tese de Doutorado. Universidade Federal de Minas Gerais. Doutorado em Ciência da Informação. 2016. Disponível em: https://repositorio.ufmg.br/bitstream/1843/BUOS-AWCNP9/1/ppgcienciainformacao_danielmendesbarbosa_tesedoutorado.pdf. Acesso em: 19 out. 2022.

BARBOSA, D. M.; BAX, M. P. Fundamentações teóricas para a criação de um modelo de gestão da informação para o contexto da avaliação de cursos de graduação. **Biblionline**, v. 12, n. 2, p. 37-52. 2016. Disponível em: <https://periodicos.ufpb.br/index.php/biblio/article/view/27890>. Acesso em: 10 abr. 2022.

BARBOSA, R. A. **A assistência ao estudante da residência universitária da UFPB**. Dissertação de Mestrado. Universidade Federal da Paraíba. Mestrado em Serviço Social. 2009. Disponível em: <https://repositorio.ufpb.br/jspui/handle/tede/7212>. Acesso em: 10 mai. 2022.

BARBOSA, R. R. Gestão da informação e do conhecimento: origens, polêmicas e perspectivas. **Informação & Informação**, v. 13, n. 1, p. 1-25, 2008. Disponível em: <https://brapci.inf.br/index.php/res/v/33432>. Acesso em: 18 out. 2022.

BARBOSA, R. R. Gestão da informação e gestão do conhecimento: evolução e conexões. **Perspectivas em Ciência da Informação**, p. 168-186, 2020. Disponível em: <https://periodicos.ufmg.br/index.php/pci/article/view/22287>. Acesso em: 24 out. 2022.

BARRETO, A. A. Uma história da Ciência da Informação. In: TOUTAIN, L. M. B. B. **Para entender a ciência da informação**. Salvador: EDUFBA, 2012. Disponível em: <https://repositorio.ufba.br/bitstream/ufba/145/1/Para%20entender%20a%20ciencia%20da%20informacao.pdf>. Acesso em: 10 mai. 2022.

BARROS, C. S. O. **Influência do plano nacional de assistência estudantil sobre indicadores de desempenho acadêmico, evasão e retenção dos discentes cotistas no período da pandemia de Covid-19 na UFPB**. Dissertação de Mestrado. Universidade Federal da Paraíba. Mestrado Profissional em Gestão nas Organizações Aprendentes. 2022. Disponível em: <https://repositorio.ufpb.br/jspui/handle/123456789/24517>. Acesso em: 30 mar. 2022.

BARROS, D. S. **Dimensões metacognitivas no comportamento de busca de informação: Estudo do usuário no Arquivo Público do Estado do Maranhão**. Dissertação de Mestrado. Universidade Federal da Paraíba. Mestrado em Ciência da Informação. 2008. Disponível em: <http://www.referenciasarquivisticas.fci.unb.br:8080/jspui/bitstream/123456789/83/1/dissertacao%20Dirlene%20Santos%20Barros.pdf>. Acesso em: 28 out. 2022.

BATISTA, G. E. A. P. **Pré-processamento de dados em aprendizado de máquina supervisionado**. Tese de Doutorado. Universidade de São Paulo. Doutorado em Ciências - Ciência da Computação e Matemática Computacional. 2003. Disponível em: <https://teses.usp.br/teses/disponiveis/55/55134/tde-06102003-160219/publico/TeseDoutorado.pdf>. Acesso em: 24 ago. 2022.

BATISTA, N A. et al. Dados de Múltiplas Fontes da Web: coleta, integração e pré-processamento. **XXIV Simpósio Brasileiro de Sistemas Multimídia e Web**, Sociedade

Brasileira de Computação, 2018. Disponível em: <https://sol.sbc.org.br/livros/index.php/sbc/catalog/view/8/19/58-1>. Acesso em: 25 ago. 2022.

BAPTISTA, T. W. F.; REZENDE, M. **A ideia de ciclo na análise de políticas públicas**. In: MATTOS, R. A.; BAPTISTA, T. W. F. (org). Caminhos para análise das políticas de saúde. Porto Alegre: Rede Unida, 2015.

BAX, M. P.; DIAS, E. W. A abordagem “Construção de Sentido”. **Workshop da ECA/USP**, São Paulo, 1997. Disponível em: https://www.researchgate.net/profile/Marcello-Bax/publication/299435657_Relatorio_A_abordagem_Construcao_de_Sentido_apresentada_em_Workshop_da_ECAUSP_em_Sao_Paulo/inks/56f69dbd08ae95e8b6d2b95f/Relatorio-A-abordagem-Construcao-de-Sentido-apresentada-em-Workshop-da-ECA-USP-em-Sao-Paulo.pdf. Acesso em: 27 out. 2022.

BERGERON, P. Information resources management. **Annual Review of Information Science and Technology (ARIST)**, v. 31, p. 263-300, 1996. Disponível em: <https://www.learntechlib.org/p/82517/>. Acesso em: 18 out. 2022.

BETZEK, S. B. F. **Avaliação do Programa Nacional de Assistência Estudantil-PNAES na UTFPR Câmpus Medianeira**. Dissertação de Mestrado. UNESP. Mestrado em Educação. 2015. Disponível em: <https://repositorio.unesp.br/handle/11449/124477>. Acesso em: 31 out. 2022.

BIANCHI, L.; ADAMCZYK, W. Políticas nacionais de assistência estudantil: síntese de evidências sobre programas de permanência. **Evidências Express**, v. 1, p. 2021. Disponível em: <https://repositorio.enap.gov.br/handle/1/7050>. Acesso em: 25 jun. 2022.

BIAZUS, C. A. **Sistema de fatores que influenciam o aluno a evadir-se dos cursos de graduação na UFSM e na UFSC: um estudo nos cursos de Ciências Contábeis**. Tese de Doutorado. Universidade Federal de Santa Catarina. Doutorado em Engenharia de Produção. 2004. Disponível em: <https://repositorio.ufsc.br/handle/123456789/87138>. Acesso em: 30 jul. 2022.

BIEMBENGUT, T. M.; PACHECO, P. S.; CONINCK, J. C. P. Análise do desempenho acadêmico de cotistas e não cotistas na UTFPR-Curitiba. **Cadernos de Pesquisa**, p. 35-48, 2018. Disponível em: <https://periodicoseletronicos.ufma.br/index.php/cadernosdepesquisa/article/view/8756>. Acesso em: 20 jul. 2022.

BITTAR, O. J. et al. Sistemas de informação em saúde e sua complexidade. **Revista de Administração em Saúde**, v. 18, n. 70, 2018. Disponível em: <https://cqh.org.br/ojs-2.4.8/index.php/ras/article/view/77>. Acesso em: 20 mar. 2023.

BITTENCOURT, L. et al. The internet of things, fog and cloud continuum: Integration and challenges. **Internet of Things**, v. 3, p. 134-155, 2018. Disponível em: <https://doi.org/10.1016/j.iot.2018.09.005>. Acesso em: 20 ago. 2022.

BITTENCOURT, M. F. N.; FERREIRA, P. A.; BRITO, M. J. Avaliação do processo de implementação de obras públicas em universidades federais: um estudo do programa Reuni. **Revista Gestão Universitária na América Latina-GUAL**, v. 10, n. 1, p. 79-102, 2017. Disponível em: <https://www.redalyc.org/journal/3193/319349826005/>. Acesso em: 8 jul. 2022.

BOCCATO, V. R. C.; FERREIRA, E. M. Estudo comparativo entre o grupo focal e o protocolo verbal em grupo no aprimoramento de vocabulário controlado em fisioterapia: uma proposta metodológica qualitativa-cognitiva. **Revista de Ciência da Informação e Documentação**,

[S. I.], v. 5, n. 1, p. 47-68, 2014. Disponível em: <https://www.revistas.usp.br/incid/article/view/63986>. Acesso em: 25 jan. 2024.

BOLSONI-SILVA, A. T. et al. Prediction of course completion by students of a university in Brazil. **Psico-USF**, v. 23, p. 425-436, 2018. Disponível em: <https://www.scielo.br/ijpusf/a/9MkPLfFq9GHyf5bJjJpXMZ4F/?format=pdf&lang=en>. Acesso em: 29 abr. 2022.

BORST, W. N. **Construction of Engineering Ontologies for Knowledge Sharing and Reuse**. PhD thesis, University of Twente, Centre for Telematics and Information Technology. 1997. Disponível em: <https://research.utwente.nl/en/publications/construction-of-engineering-ontologies-for-knowledge-sharing-and->. Acesso em: 21 set. 2022.

BRAGA, G. P. O.; PRÁ, K. R. Assistência Estudantil: delimitações históricas e o novo quadro da conjuntura brasileira. **Cadernos Cajuína**, v. 6, n. 3, p. 5-20, 2021. Disponível em: <https://cadernoscajuina.pro.br/revistas/index.php/cadcajuina/article/view/489>. Acesso em: 10 out. 2022.

BRANDT, J. Z. **Ensino superior no Brasil: fatores influenciadores no desempenho acadêmico dos alunos de graduação**. Dissertação de Mestrado. Universidade do Minho. Mestrado em Administração Pública. 2017. Disponível em: <https://repositorium.sdum.uminho.pt/handle/1822/49271>. Acesso em: 03 ago. 2022.

BRASIL. Decreto Nº 6.096 de 24 de abril de 2007. **Institui o Programa de Apoio a Planos de Reestruturação e Expansão das Universidades Federais - REUNI**. Brasília, Brasil. 2007a. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2007/decreto/d6096.htm. Acesso em: 15 abr. 2022.

BRASIL. Portaria Normativa Nº 39, de 12 de dezembro de 2007. **Institui o Programa Nacional de Assistência Estudantil - PNAES**. Brasília, Brasil. 2007b. Disponível em: http://portal.mec.gov.br/arquivos/pdf/portaria_pnaes.pdf. Acesso em: 2 jul. 2022.

BRASIL. Decreto Nº 7.234, de 19 de julho de 2010. **Dispõe sobre o Programa Nacional de Assistência Estudantil (PNAES)**. Brasília, Brasil. 2010. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2010/decreto/d7234.htm. Acesso em: 3 mar. 2022.

BRASIL. Ministério da Educação. **Análise sobre a expansão das Universidades Federais 2003 a 2012**. Brasília, Brasil. 2012a. Disponível em: <http://portal.mec.gov.br/docman/janeiro-2013-pdf/12386-analise-expansao-universidade-federais-2003-2012-pdf>. Acesso em: 13 set. 2022.

BRASIL. Lei Nº 12.711, de 29 de agosto de 2012. **Dispõe sobre o ingresso nas universidades federais e nas instituições federais de ensino técnico de nível médio e dá outras providências**. Brasília, Brasil. 2012b. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2010/decreto/d7234.htm. Acesso em: 3 mar. 2022.

BRASIL. Ministério da Educação. **Plataforma Nilo Peçanha**. Brasília: MEC/SETEC, 2024. Disponível em: <https://www.gov.br/mec/pt-br/pnp>. Acesso em: 21 dez. 2024.

BREITMAN, K. K.; FELICÍSSIMO, C. H.; CASANOVA, M. A. CATO-A Lightweight Ontology Alignment Tool. **CAiSE Short Paper Proceedings**. 2005. Disponível em:

<http://www-di.inf.puc-rio.br/~casanova/Publications/Papers/2005-Papers/2005-CAISE-Breitman.pdf>. Acesso em: 26 set. 2022.

BRITO, A. C. T.; COSTA, A. K. T.; ALMEIDA, V. L. Gestão universitária: políticas comparadas de monitoramento do programa bolsa permanência em universidades federais brasileiras. **Revista Ibero-Americana de Estudos em Educação**, p. 1862-1875, 2019. Disponível em: <https://periodicos.fclar.unesp.br/iberoamericana/article/view/13036>. Acesso em: 12 abr. 2022.

BRITTO, G.; RUY, F. B.; AZEVEDO, C. L. B. Um ambiente para integração de dados abertos relativos à despesa pública. **Ontobras**, 13, 2020. Disponível em: <https://ceur-ws.org/Vol-2728/paper13.pdf>. Acesso em: 21 mar. 2023.

BROWN, M. G. **Keeping score**: Using the right metrics to drive world-class performance. Productivity Press, 2020.

BROWN, S. D. et al. Social cognitive predictors of college students' academic performance and persistence: A meta-analytic path analysis. **Journal of Vocational Behavior**, v. 72, n. 3, p. 298-308, 2008. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0001879107001182>. Acesso em: 27 jul. 2022.

BUENO, J. L. O. A evasão de alunos. **Paidéia**, p. 9-16, 1993. Disponível em: <https://www.scielo.br/j/paideia/a/6F8TQQbf5N3ZsDPGzJJXj9p/?lang=pt>. Acesso em: 10 mai. 2022.

CABRAL, D. S. et al. **Problemas e desafios na integração de sistemas e dados entre diferentes esferas governamentais**: um estudo de caso do sistema Trakcare®. Dissertação de Mestrado. Universidade Católica de Brasília. Mestrado em Governança, Tecnologia e Inovação. 2021. Disponível em: <https://bdt.d.ucb.br:8443/jspui/handle/tede/2831>. Acesso em: 15 out. 2022.

CALIATTO, S. G.; ALMEIDA, L. S. Aprendizagem e rendimento acadêmico no Ensino Superior. **Revista Ibero-Americana de Estudos em Educação**, p. 1855-1876, 2020. Disponível em: <https://periodicos.fclar.unesp.br/iberoamericana/article/view/12670>. Acesso em: 01 ago. 2022.

CALIENDO, M.; KOPEINIG, S. Some practical guidance for the implementation of Propensity Score Matching. **Journal of Economic Surveys**, v. 22, n. 1, p. 31-72, 2008. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-6419.2007.00527.x>. Acesso em: 2 abr. 2024.

CAMARGO, A. M. M.; ARAÚJO, I. M. Expansão e interiorização das universidades federais no período de 2003 a 2014: perspectivas governamentais em debate. **Acta Scientiarum. Education**, v. 40, n. 1, p. e37659, 2018. Disponível em: <https://periodicos.uem.br/ojs/index.php/ActaSciEduc/article/view/37659>. Acesso em: 02 set. 2022.

CAMPOS, J. G. et al. Estado da arte sobre arquiteturas de sistemas para integração de dados. **Rio Oil & Gas Expo and Conference**, 2020. Disponível em: <https://biblioteca.ibp.org.br/rioilegas/pt-BR/search/39610?exp=>. Acesso em: 11 mai. 2022.

CAMPOS, P. M. C. **Designing a Network of Reference Ontologies for the Integration of Water Quality Data**. Dissertação de Mestrado. Universidade Federal do Espírito Santo. Mestrado em Informática. 2019. Disponível em:

https://nemo.inf.ufes.br/wp-content/papercite-data/pdf/designing_a_network_of_reference_ontologies_for_the_integration_of_water_quality_data_2019.pdf. Acesso em: 28 jul. 2022.

CAÑAVERAL, I. C. P.; SÁ, T. A. O. REUNI: Expansão, segmentação e a determinação institucional do abandono: estudo de caso na Unifal-MG. **EccoS–Revista Científica**, n. 44, p. 93-115, 2017. Disponível em: <https://periodicos.uninove.br/eccos/article/view/7899>. Acesso em: 1 mai. 2022.

CARCHEDI, L. C. et al. Onto4LA: uma ontologia para integração de dados educacionais. **Congresso Brasileiro de Informática na Educação**, p. 439, 2018. Disponível em: <http://ojs.sector3.com.br/index.php/wcbie/article/view/8269>. Acesso em: 04 jan. 2023.

CARMO, P. N. S. et al. Aprendizado de Máquina em Ações de Controle no Tribunal de Contas do Estado do Maranhão. **VII Jornada de Informática do Maranhão**, 2018. Disponível em: <http://sistemas.deinf.ufma.br/anaisjim/artigos/2018/201809.pdf>. Acesso em: 27 ago. 2022.

CARVALHO, V. A.; ALMEIDA, J. P. A.; GUIZZARDI, G. Using reference domain ontologies to define the real-world semantics of domain-specific languages. **Advanced Information Systems Engineering**, CAiSE 2014, Lecture Notes in Computer Science, v. 8484, 2014. Disponível em: https://doi.org/10.1007/978-3-319-07881-6_33. Acesso em: 25 set. 2022.

CASIRAGHI, B. et al. Rendimento acadêmico no Ensino Superior: variáveis pessoais e socioculturais do estudante. **Revista Praxis**, v. 12, n. 24, 2020. Disponível em: <https://revistas.unifoa.edu.br/praxis/article/view/3373>. Acesso em: 30 jul. 2022.

CASTELLS, M. **A sociedade em rede**. São Paulo: Paz e Terra, 2011.

CASTELO BRANCO, U. V.; NAKAMURA, P. H.; JEZINE, E. Permanência na educação superior no Brasil: construção de uma escala de medida. **Revista de la Asociación de Sociología de la Educación (RASE)**, v. 10, 2017. Disponível em: <https://roderic.uv.es/handle/10550/65147>. Acesso em: 03 mai. 2022.

CASTRO, A. K. S.; TEIXEIRA, M. A. P. A evasão em um curso de psicologia: uma análise qualitativa. **Psicologia em Estudo**, v. 18, n. 2, p. 199-209, 2013. Disponível em: <https://www.scielo.br/j/pe/a/YRM6f59TQX5stXQhyNK459c/?lang=pt>. Acesso em: 10 abr. 2022.

CASTRO, F. F.; SIMIONATO, A. C. Revisitando ontologia e metadados à luz dos ambientes informacionais digitais. **Perspectivas em Ciência da Informação**, v. 25, p. 3-23, 2021. Disponível em: <https://www.scielo.br/j/pci/a/nr5YfqhgBbrjJHyHJhyYLc/?format=html&stop=previous&lang=pt>. Acesso em: 21 mar. 2023.

CASTRO, T. L. **Evasão nos cursos de licenciatura do instituto federal de educação, ciência e tecnologia de Minas Gerais**. Tese de Doutorado. Universidade Federal de Minas Gerais (UFMG). Doutorado em Educação. 2019. Disponível em: <http://hdl.handle.net/1843/30542>. Acesso em: 01 mai. 2022.

CAVALCANTI, I. T. N. **Análise dos diferenciais de desempenho dos estudantes cotistas e não cotistas da UFBA pelo propensity score matching**. 2015. Dissertação (Mestrado em Economia). Universidade Federal da Bahia. 2015. Disponível em: <https://repositorio.ufba.br/handle/ri/18125>. Acesso em: 11 abr. 2024.

CERATTI, M. R. N. **Evasão escolar**: causas e consequências. Curitiba/PR, 2008. Disponível em:

<https://www.seduc.go.gov.br/imprensa/documentos/Arquivos/15%20-%20Manual%20de%20Gest%C3%A3o%20Pedag%C3%B3gico%20e%20Administrativo/2.10%20Combate%20%C3%A0%20evas%C3%A3o/EVAS%C3%83O%20ESCOLAR%20-%20CAUSAS%20E%20CONSEQUENCIAS.pdf>. Acesso em: 16 jun. 2022.

CERRI, R.; CARVALHO, A. C. P. Aprendizado de máquina: breve introdução e aplicações. **Cadernos de Ciência & Tecnologia**, v. 34, n. 3, p. 297-313, 2017. Disponível em: <https://ainfo.cnptia.embrapa.br/digital/bitstream/item/184785/1/Aprendizado-de-maquina-breve-introducao.pdf>. Acesso em: 29 ago. 2022.

CESPEDES, J. G. et al. Avaliação de impacto do Programa de Permanência Estudantil da Universidade Federal de São Paulo. **Ensaio: Avaliação e Políticas Públicas em Educação**, v. 29, p. 1067-1091, 2021. Disponível em: <https://www.scielo.br/i/enso/a/Jbgmjrb7dTKdFKGHvVPWNC/abstract/?lang=pt>. Acesso em: 31 out. 2022.

CHEN, D. Walmart sales prediction based on random forest model and application of feature importance. **Applied and Computational Engineering**, v. 53, p. 264-273, 2024. Disponível em: <https://www.ewadirect.com/proceedings/ace/article/view/11028>. Acesso em: 15 abr. 2025.

CHENG, G; DU, Q.; MA, H. The design and implementation of ontology and rules based knowledge base for transportation. **2008 International Conference on Computer Science and Software Engineering (IEEE)**, 2008. p. 1035-1038. Disponível em: <https://ieeexplore.ieee.org/abstract/document/4722519>. Acesso em: 24 jul. 2022.

CHOI, T. M.; WALLACE, S. W.; WANG, Y. Big data analytics in operations management. **Production and Operations Management**, v. 27, n. 10, p. 1868-1883, 2018. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/poms.12838>. Acesso em: 24 ago. 2022.

CHOO, C. W. Information Management for the Intelligent Organization: roles and implications for the information professions. **Digital Libraries Conference**, 1995. Disponível em: <http://choo.ischool.utoronto.ca/FIS/respub/DLC95.pdf>. Acesso em: 18 out. 2022.

CHOO, C. W. Closing the cognitive gaps: how people process information. **Financial Times of London**, v. 22, 1999. Disponível em: <http://choo.ischool.utoronto.ca/FIS/respub/FThis/default.html>. Acesso em: 29 out. 2022.

CHOO, C. W. **Information Management for the Intelligent Organization**: the art of scanning the environment. Medford: Information Today, 2002.

CHOO, C. W. **A organização do conhecimento**: como as organizações usam a informação para criar significado, construir conhecimento e tomar decisões. São Paulo: SENAC, 2006.

CIANCONI, R. B. **Gestão do conhecimento**: visão de indivíduos e organizações no Brasil. Tese de Doutorado. Universidade Federal do Rio de Janeiro. Doutorado em Ciência da Informação. 2003. Disponível em: <http://ridi.ibict.br/handle/123456789/675>. Acesso em: 23 out. 2022.

CISLAGHI, R. **Um modelo de sistema de gestão do conhecimento em um framework para a promoção da permanência discente no ensino de graduação**. Tese de Doutorado. Universidade Federal de Santa Catarina. Doutorado em Engenharia e Gestão do Conhecimento. 2008. Disponível em:

<https://repositorio.ufsc.br/bitstream/handle/123456789/91544/250753.pdf?sequence>. Acesso em: 11 abr. 2022.

CISLAGHI, J. F.; SILVA, M. T. O plano nacional de assistência estudantil e o Reuni: ampliação de vagas versus garantia de permanência. **SER Social**, v. 14, n. 31, p. 489-512, jul-dez 2012. Disponível em: https://periodicos.unb.br/index.php/SER_Social/article/view/13025. Acesso em: 2 jul. 2022.

COIMBRA, C. L.; SILVA, L. B.; COSTA, N. C. D. A evasão na educação superior: definições e trajetórias. **Educação e Pesquisa**, v. 47, p. e228764, 2021. Disponível em: <https://www.redalyc.org/journal/298/29866573016/html/>. Acesso em: 30 jul. 2022.

COMISSÃO ESPECIAL DE ESTUDOS SOBRE A EVASÃO NAS UNIVERSIDADES PÚBLICAS BRASILEIRAS. **Diplomação, retenção e evasão nos cursos de graduação em instituições de ensino superior públicas**. Brasília, DF, 1996. Disponível em: https://www.andifes.org.br/wp-content/files_flutter/Diplomacao_Retencao_Evasao_Graduacao_em_IES_Publicas-1996.pdf. Acesso em: 19 abr. 2022.

CONCEIÇÃO, A. P.; LONGHINI, T. M.; OLIVEIRA, Y. A. Evasão em curso de Engenharia de Produção de um Instituto Federal. **Revista Latino-Americana de Inovação e Engenharia de Produção**, v. 8, n. 13, p. 121-141, 2020. Disponível em: <https://revistas.ufpr.br/relainep/article/view/73030>. Acesso em: 10 ago. 2022.

CORDEIRO, R. A.; SILVA, A. B. Os estilos de aprendizagem influenciam o desempenho acadêmico dos estudantes de finanças?. **Revista de Administração da Universidade Federal de Santa Maria**, v. 5, n. 2, p. 243-261, 2012. Disponível em: <https://www.redalyc.org/pdf/2734/273424461005.pdf>. Acesso em: 16 jul. 2022.

COSTA, E. et al. Mineração de dados educacionais: conceitos, técnicas, ferramentas e aplicações. **Jornada de Atualização em Informática na Educação**, v. 1, n. 1, p. 1-29, 2013. Disponível em: <http://ojs.sector3.com.br/index.php/pie/article/view/2341>. Acesso em: 25 ago. 2022.

COSTA, L. F.; SILVA, A. C. P.; RAMALHO, F. A. (Re)visitando os estudos de usuário: entre a “tradição” e o “alternativo”. **DataGramZero**, v. 10, n. 4, p. 1-12, 2009. Disponível em: <https://brapci.inf.br/index.php/res/v/6946>. Acesso em: 28 out. 2022.

COSTA, S. D. **Apoio à decisão na gestão de pessoas em projetos de software: uma abordagem utilizando simulação com dinâmica de sistemas**. Dissertação de Mestrado. Universidade Federal de Viçosa. Mestrado em Ciência da Computação. 2012. Disponível em: <https://www.locus.ufv.br/handle/123456789/2645>. Acesso em: 20 jul. 2022.

CRESWELL, J. W. **Research Design: Qualitative, Quantitative, and Mixed Methods Approachs**. SAGE, 2014.

CRUZ, M. M. L. et al. Semanticsus: Um portal semântico baseado em ontologias e dados interligados para acesso, integração e visualização de dados do SUS. **Anais Estendidos do XIX Simpósio Brasileiro de Computação Aplicada à Saúde**. SBC, 2019. p. 13-18. Disponível em: https://sol.sbc.org.br/index.php/sbcas_estendido/article/view/6277. Acesso em: 20 mar. 2023.

CUNHA, C. G. S. **Avaliação de políticas públicas e programas governamentais: tendências recentes e experiências no Brasil**. Revista Estudos de Planejamento, n. 12, 2018. Disponível em: <http://200.198.145.164/index.php/estudos-planejamento/article/view/4298>. Acesso em: 10 jun. 2023.

DAHER JUNIOR, F. J.; BORGES, J. Ciência da Informação: uma utopia transdisciplinar?. **AtoZ: novas práticas em informação e conhecimento**, v. 10, n. 2, p. 89-96, 2021. Disponível em: <https://revistas.ufpr.br/atoz/article/view/79257/43724>. Acesso em: 02 jan. 2023.

DAVENPORT, T. H. **Dominando a Gestão da Informação**. Porto Alegre, Bookman, 2004.

DAVENPORT, T. H. **Ecologia da informação**. São Paulo: Futura, 2002.

DAVENPORT, T. H.; PRUSAK, L. **Conhecimento empresarial: como as organizações gerenciam o seu capital intelectual**. Rio de Janeiro: Campus, 1998.

DAVOK, D. F.; BERNARD, R. P. Avaliação dos índices de evasão nos cursos de graduação da Universidade do Estado de Santa Catarina-UDESC. **Revista da Avaliação da Educação Superior (Campinas)**, v. 21, p. 503-522, 2016. Disponível em: <https://www.scielo.br/j/aval/a/5VJRg7PrXDTQ5mYXK95rh8r/?format=html&lang=pt>. Acesso em: 05 mai. 2022.

DEGBELO, A. A snapshot of ontology evaluation criteria and strategies. **Proceedings of the 13th International Conference on Semantic Systems**, p. 1-8, 2017. Disponível em: <https://dl.acm.org/doi/pdf/10.1145/3132218.3132219>. Acesso em: 24 jan. 2024.

DERVIN, B. An Overview of Sense-Making: concepts, methods and results to date. **Annual Meeting of the International Communications Association**. 1983. Disponível em: <https://faculty.washington.edu/wpratt/MEBI598/Methods/An%20Overview%20of%20Sense-Making%20Research%201983a.htm>. Acesso em: 26 fev. 2023.

DETLOR, B. Information Management. **International Journal of Information Management**, v. 30, p. 103-108, 2010. Disponível em: <https://dl.acm.org/doi/10.1016/j.ijinfomgt.2009.12.001>. Acesso em: 30 nov. 2022.

DETONI, A. A. et al. Uma Ontologia de Referência para Autorização Orçamentária e Execução da Despesa Pública. **iSys-Brazilian Journal of Information Systems**, v. 11, n. 3, p. 4-53, 2018. Disponível em: <https://sol.sbc.org.br/journals/index.php/isys/article/view/369>. Acesso em: 19 nov. 2022.

DEVEDZIC, V. Understanding ontological engineering. **Communications of the ACM**, v. 45, n. 4, 2002. Disponível em: <https://dl.acm.org/doi/abs/10.1145/505248.506002>. Acesso em: 23 set. 2022.

DE GIACOMO, G. et al. Using Ontologies for Semantic Data Integration. In: Flesca, S., Greco, S., Masciari, E., Saccà, D. (eds) A Comprehensive Guide Through the Italian Database Research Over the Last 25 Years. **Studies in Big Data**, Springer, v. 31, 2018. Disponível em: https://doi.org/10.1007/978-3-319-61893-7_11. Acesso em: 22 abr. 2022.

DIAS, M. M. K.; BELLUZZO, R. C. B. **Gestão da informação em ciência e tecnologia sob a ótica do cliente**. Bauru: EDUSC, 2003.

DIAS, M. S. L.; CARVALHO, E. R. A assistência estudantil no IFPR: a construção de um objeto de estudo. **Cadernos Cajuína**, v. 6, n. 3, p. 21-35, 2021. Disponível em: <https://cadernoscajuina.pro.br/revistas/index.php/cadcajuina/article/view/490>. Acesso em: 13 jul. 2022.

DIAS, S. M. B.; COSTA, S. L. A permanência no ensino superior e as estratégias institucionais de enfrentamento da evasão. **Jornal de políticas educacionais**, v. 9, n. 17/18, 2016. Disponível em: <https://revistas.ufpr.br/jpe/article/view/38650>. Acesso em: 30 jun. 2022.

DIAZ, E.; WATTS, M. Metrics-driven decision-making improves performance at a complex process facility. **Process Safety Progress**, v. 39, n. 2, p. e12092, 2020. Disponível em: <https://aiche.onlinelibrary.wiley.com/doi/abs/10.1002/prs.12092>. Acesso em: 22 out. 2022.

DING, Y.; FOO, S. Ontology research and development. Part 1 - a review of ontology generation. **Journal of Information Science**, v. 28, n. 2, 123–136, 2022. Disponível em: <https://doi.org/10.1177/016555150202800204>. Acesso em: 26 set. 2022.

DRAIBE, S. M. **Avaliação de implementação**: esboço de uma metodologia de trabalho em políticas públicas. In: BARREIRA, M. C.; CARVALHO, M. C. (org). Tendências e perspectivas na avaliação de políticas e programas sociais. São Paulo : IEE/PUC-SP, 2001. Disponível em: <http://bibliotecadigital.abong.org.br/handle/11465/1763>. Acesso em: 14 mai. 2024.

ECKERSON, W. W. **Performance dashboards**: measuring, monitoring, and managing your business. John Wiley & Sons, 2010.

ELLIS, D. A behavioural approach to information retrieval system design. **Journal of Documentation**, v. 45, n. 3, p. 171-212. 1989. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/eb026843/full/html>. Acesso em: 21 mar. 2023.

ELOI, S. S. S. et al. O Programa Nacional de Assistência Estudantil–PNAES na Universidade Federal de Viçosa (UFV): uma avaliação na percepção dos beneficiários. **Revista Gestão Universitária na América Latina-GUAL**, v. 12, n. 3, p. 106-129, 2019. Disponível em: <https://periodicos.ufsc.br/index.php/gual/article/view/1983-4535.2019v12n3p106>. Acesso em: 19 jun. 2022.

FACCIONI FILHO, M. **Internet das coisas**. Palhoça, Unisul Virtual, 2016.

FACELI, K. et al. **Inteligência Artificial**: uma abordagem de Aprendizado de Máquina. Rio de Janeiro: LTC, 2011.

FAGUNDES, C. V.; LUCE, M. B.; ESPINAR, S. R. O desempenho acadêmico como indicador de qualidade da transição Ensino Médio-Educação Superior. **Ensaio: avaliação e políticas públicas em educação**, v. 22, p. 635-669, 2014. Disponível em: <https://www.scielo.br/j/ensaio/a/pmMZfDzDfshjDK7G8XDLPdC>. Acesso em: 17 jul. 2022.

FALBO, R. A. SABiO: Systematic Approach for Building Ontologies. **ONTO. COM/ODISE@FOIS**. 2014. Disponível em: http://ceur-ws.org/Vol-1301/ontocomodise2014_2.pdf. Acesso em: 15 ago. 2022.

FARIAS, D. G. **A política nacional de assistência estudantil nas universidades públicas federais**: os desafios ao trabalho profissional dos/as assistentes sociais no âmbito de programas e acessos estudantis. Dissertação de Mestrado. Universidade de Brasília. Mestrado em Política Social. 2021. Disponível em: <https://repositorio.unb.br/handle/10482/41724>. Acesso em: 27 set. 2022.

FARINELLI, F.; ALMEIDA, M. B.; MELO, S. O papel das ontologias na interoperabilidade de sistemas de informação: reflexões na esfera governamental. **XIV Encontro Nacional de Pesquisa em Ciência da Informação**, 2013. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/185252>. Acesso em: 21 mar. 2023.

FAVA, H. L.; HALL, R. J.; CINTRA, R. F. Indicadores na Assistência Estudantil: o que a literatura “diz”? **Jornal de Políticas Educacionais**, v. 16, n. 1, 2022. Disponível em: <https://revistas.ufpr.br/jpe/article/view/83591>. Acesso em: 14 fev. 2023.

FERREIRA, A.; CRISÓSTOMO, J. A influência do desempenho acadêmico na carreira profissional: um estudo de caso em um curso de engenharia. **Revista de ensino de engenharia**, v. 30, n. 1, p. 35-44, 2012. Disponível em: <http://revista.educacao.ws/revista/index.php/abenge/article/view/96>. Acesso em: 20 jul. 2022.

FERREIRA, F. B. et al. Determinantes do rendimento acadêmico dos discentes dos cursos de agronomia e zootecnia em uma instituição de ensino superior privada. **Research, Society and Development**, v. 10, n. 2, p. e43310212747-e43310212747, 2021. Disponível em: <https://rsdjournal.org/index.php/rsd/article/view/12747>. Acesso em: 27 jul. 2022.

FERREIRA, F. M. **A atuação da UFU frente a pandemia de covid-19—a gestão de mudanças em cenários de contingências**. Dissertação de Mestrado. Universidade Federal de Uberlândia. Mestrado em Gestão Organizacional. 2021. Disponível em: <https://repositorio.ufu.br/handle/123456789/33961>. Acesso em: 20 dez. 2022.

FERREIRA, M. A. **Determinantes do desempenho discente no ENADE em cursos de Ciências Contábeis**. Dissertação de Mestrado. Universidade Federal de Uberlândia. Mestrado em Ciências Contábeis, 2015. Disponível em: <http://repositorio.ufu.br/handle/123456789/12620>. Acesso em: 10 jul. 2022.

FIALHO, M. G. D.; PRESTES, E. M. T. Evasão escolar no curso de pedagogia da UFPB: na compreensão dos gestores educacionais. **Mpgoa**, João Pessoa, v. 3, n. 1, p. 42-63, 2014. Disponível em: <https://doi.org/10.23179/2317-9082>. Acesso em: 12 dez. 2022.

FIGUEIREDO, N. G. S.; SALLES, D. M. R. Educação profissional e evasão escolar em contexto: motivos e reflexões. **Ensaio: Avaliação e Políticas Públicas em Educação**, v. 25, p. 356-392, 2017. Disponível em: <https://www.scielo.br/j/ensaio/a/Bw8WKpzdP3w8qn5zL68C3sq/abstract/?lang=pt>. Acesso em: 07 fev. 2022.

FINATTI, B. E.; ALVES, J. M.; SILVEIRA, R. J. Perfil Sócio, Econômico e Cultural dos Estudantes da Universidade Estadual de Londrina-UEL—Indicadores para Implantação de uma Política de Assistência Estudantil. **Libertas**, v. 6, n. 1/2, 2006. Disponível em: <https://periodicos.ufjf.br/index.php/libertas/article/view/18155>. Acesso em: 3 jul. 2022.

FISHER, D. H.; PAZZANI, M. J.; LANGLEY, P. **Concept formation: Knowledge and experience in unsupervised learning**. Morgan Kaufmann, 2014.

FLORES, J. T. **A política de assistência estudantil na Universidade Federal do Rio Grande do Sul: um estudo de caso para analisar a eficácia das ações através da taxa de desperdício de matrícula**. Dissertação de Mestrado. Universidade Federal do Rio Grande do Sul. Mestrado em Economia. 2022. Disponível em: <https://www.lume.ufrgs.br/handle/10183/236383>. Acesso em: 10 jul. 2022.

FONAPRACE. Fórum Nacional de Pró-Reitores de Assuntos Comunitários e Estudantis. **Revista Comemorativa: 25 anos: histórias, memórias e múltiplos olhares**. 2012. Disponível em: http://www.proae.ufu.br/sites/proae.ufu.br/files/media/arquivo/revista_fonaprace_25_anos.pdf. Acesso em: 2 jul. 2022.

FONSECA, C. M. et al. Ontology-Driven Conceptual Modeling as a Service. **JOWO**. 2021. Disponível em: <https://www.academia.edu/download/81075826/paper29-FOMI.pdf>. Acesso em: 01 set. 2023.

FONSECA, L. B.; AZEVEDO, C. L.; ALMEIDA, J. P. A. Mapeando dados governamentais com uma ontologia de organizações. **LOD Brasil**, 2014. Disponível em: https://www.researchgate.net/profile/Lucas-Fonseca-15/publication/269109674_Mapeando_Dados_Governamentais_com_uma_Ontologia_de_Organizacoes/links/5481a6d30cf22525dc_b62a2a/Mapeando-Dados-Governamentais-com-uma-Ontologia-de-Organizacoes.pdf. Acesso em: 10 fev. 2023.

FRANÇA, M. T. A.; GONÇALVES, F. O. Provisão pública e privada de educação fundamental: Diferenças de qualidade medidas por meio de propensity score. **Economia Aplicada**, v. 14, p. 373-390, 2010. Disponível em: <https://www.scielo.br/j/ecoa/a/yPH7xsYHRP4N9jwSRPhfpBJ/>. Acesso em: 9 abr. 2024.

FREITAS, F. C. L. **Avaliação do Programa Nacional de Assistência Estudantil na UFC: a permanência discente como direito à educação dos estudantes do Campus de Sobral**. Dissertação de Mestrado. Universidade Federal do Ceará. Mestrado em Avaliação de Políticas Públicas. 2021. Disponível em: <https://repositorio.ufc.br/handle/riufc/58080>. Acesso em: 15 jan. 2023.

FREITAS, F.; MENDES, J. M. Grandes dados e os desafios atuais para a produção de conhecimento. In: Portugal, território de territórios. **Atas do IX Congresso Português de Sociologia**. Associação Portuguesa de Sociologia, 2017. Disponível em: <https://estudogeral.sib.uc.pt/handle/10316/96943>. Acesso em: 23 ago. 2022.

FRIEDMAN, C. P.; FLYNN, A. J. Computable knowledge: an imperative for learning health systems. **Learning health systems**, v. 3, n. 4, 2019. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6802532/>. Acesso em: 10 nov. 2021.

FRITSCH, R.; ROCHA, C. S.; VITELLI, R. F. A evasão nos cursos de graduação em uma instituição de ensino superior privada. **Revista Educação em Questão**, v. 52, n. 38, p. 81-108, 2015. Disponível em: <https://periodicos.ufrn.br/educacaoemquestao/article/view/7963/>. Acesso em: 20 nov. 2021.

GALHARDO, E. et al. Desempenho acadêmico e frequência dos estudantes ingressantes pelo Programa de Inclusão da UNESP. **Revista da Avaliação da Educação Superior (Campinas)**, v. 25, p. 701-723, 2020. Disponível em: <https://www.scielo.br/j/aval/a/7zWxrQGM9PJGfTfCyrYyLYr/?format=html&lang=pt>. Acesso em: 20 jul. 2022.

GAMA, B. B. O. **Determinantes da evasão universitária e impacto no gasto público**. Dissertação de Mestrado. Universidade Federal de Uberlândia. Mestrado em Gestão Organizacional. 2018. Disponível em: <http://repositorio.ufu.br/handle/123456789/22636>. Acesso em: 25 nov. 2021.

GARCIA, C. M.; ABILIO, R.; MALHEIROS, N. Abordagens e tecnologias para integração de sistemas: Um estudo de caso na Universidade Federal de Lavras. **Revista de Sistemas de Informação da FSMA**, v. 15, p. 11-22, 2015. Disponível em: https://www.researchgate.net/profile/Cristiano-Garcia-4/publication/320957389_Abordagens_e_Tecnologias_para_Integracao_de_Sistemas_Um_Estudo_de_Caso_na_Universidade_Federal_de_Lavras/links/5a046603aca272b06cadc3cd/Abordagens-e-Tecnologias-para-Integracao-de-Sistemas-Um-Estudo-de-Caso-na-Universidade-Federal-de-Lavras.pdf. Acesso em: 11 jun. 2022.

GARCÍA, E. et al. Drawbacks and solutions of applying association rule mining in learning management systems. **Proceedings of the international workshop on applying data mining in e-learning (ADML 2007)**, Crete, Greece. sn, p. 13-22, 2007. Disponível em: https://www.academia.edu/download/68536575/distance_education_a_web_usage_mining_case_study_for_the.pdf. Acesso em: 03 ago. 2022.

GASQUE, K. C. G. D.; COSTA, S. M. S. Evolução teórico-metodológica dos estudos de comportamento informacional de usuários. **Ciência da Informação**, v. 39, p. 21-32, 2010. Disponível em: <https://www.scielo.br/ci/a/wzMJ66VNkZZxxKxnk7G3ktm/abstract/?lang=pt>. Acesso em: 26 out. 2022.

GASQUE, K. C. G. D. **O pensamento reflexivo na busca e no uso da informação na comunicação científica**. Tese de Doutorado. Universidade de Brasília. Doutorado em Ciência da Informação. 2008. Disponível em: <https://repositorio.unb.br/handle/10482/1344>. Acesso em: 29 jul. 2022.

GERHARDT, T. E.; SILVEIRA, D. T. **Métodos de Pesquisa**. Porto Alegre: Editora da UFRGS, 2009.

GHELLER, A. A. et al. Fatores críticos de sucesso em projetos ERP cloud sob os aspectos processos, sistema e tecnologia no contexto empresarial brasileiro. **Revista de Gestão e Projetos**, v. 8, n. 2, p. 127-140, 2017. Disponível em: <https://periodicos.uninove.br/gep/article/view/9671>. Acesso em: 12 jul. 2022.

GIBSON, B. R.; ROGERS, T. T.; ZHU, X. Human semi-supervised learning. **Topics in cognitive science**, v. 5, n. 1, p. 132-172, 2013. Disponível em: <https://doi.org/10.1111/tops.12010>. Acesso em: 30 ago. 2022.

GOLDENBERG, M. **A arte de pesquisar: como fazer pesquisa qualitativa em Ciências Sociais**. São Paulo: Record, 2011.

GOMES, A. A. **Evasão e evadidos: o discurso dos ex-alunos sobre evasão escolar nos cursos de licenciatura**. Tese de Doutorado. Universidade Estadual Paulista. Doutorado em Educação. 1998. Disponível em: <https://repositorio.unesp.br/handle/11449/102247>. Acesso em: 10 mar. 2022.

GOMES, A. M. O.; PASSOS, G. O. A implementação do Programa Nacional de Assistência Estudantil (PNAES) nos Institutos Federais. **Revista de Políticas Públicas**, v. 22, n. 1, p. 416-442, 2018. Disponível em: <http://periodicoeletronicos.ufma.br/index.php/rppublica/article/view/9240>. Acesso em: 3 jul. 2022.

GOMES, G.; SOARES, A. B. Inteligência, habilidades sociais e expectativas acadêmicas no desempenho de estudantes universitários. **Psicologia: Reflexão e Crítica**, v. 26, p. 780-789, 2013. Disponível em: <https://www.scielo.br/j/prc/a/YMwTmfCg4gYhq4Kc8cnTJYJ/abstract/?lang=pt>. Acesso em: 20 jul. 2022.

GOMES, S. C. et al. Adoção de tecnologia da informação como estratégia para melhorar o desempenho da gestão de micro e pequenas empresas. **REMIPE-Revista de Micro e Pequenas Empresas e Empreendedorismo da Fatec Osasco**, v. 5, n. 1, p. 24-45, 2019. Disponível em: <http://remipe.fatecosasco.edu.br/index.php/remipe/article/view/165>. Acesso em: 25 out. 2022.

GOMES NETO, A.; SANTOS, A. M. A.; PRATES, T. M. Como o programa Bolsa Família impacta na educação básica? Uma análise para o Nordeste do Brasil. **Revista Econômica do Nordeste**, v. 51, n. 1, p. 51-72, 2020. Disponível em: <https://www.bnb.gov.br/revista/index.php/ren/article/download/955/804>. Acesso em: 5 abr. 2024.

GONÇALVES, I. L. **Taxa de evasão e impacto financeiro na realidade da UNIFAL-MG**. Dissertação de Mestrado. Universidade Federal de Alfenas. Mestrado em Administração Pública. 2018. Disponível em: <http://bdtd.unifal-mg.edu.br:8080/handle/tede/1162>. Acesso em: 12 jan. 2022.

GONÇALVES, J. E. A. **Método ágil de integração semântica de dados científicos baseado em ontologias**. Tese de Doutorado. Universidade Federal de Minas Gerais. Doutorado em Gestão & Organização do Conhecimento. 2020. Disponível em: <https://repositorio.ufmg.br/handle/1843/34013>. Acesso em: 12 jan. 2023.

GONÇALVES, M. Abordagem sense-making na ciência da informação: uma breve contextualização. **RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação**, v. 10, n. 1, p. 1-11, 2012. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/1906>. Acesso em: 28 out. 2022.

GONZALES JUNIOR, I. P.; SILVA, G. L. A.; SAMPAIO, L. A. A importância do sistema de informação na gestão pública como ferramenta de processo decisório. **Acta Negócios**, v. 2, n. 1, p. 9-25, 2019. Disponível em: <https://www.revistas.unasp.edu.br/actanegocios/article/download/1144/1132>. Acesso em: 01 mar. 2023.

GRUBER, T. R. Toward principles for the design of ontologies used for knowledge sharing. **International Journal Human-Computer Studies**, v. 43, n. 5-6, p. 907-928, 1995. Disponível em: <https://doi.org/10.1006/ijhc.1995.1081>. Acesso em: 26 set. 2022.

GUARINO, N. Understanding, building and using ontologies. **International journal of human-computer studies**, v. 46, n. 2-3, p. 293-310, 1997. Disponível em: <https://doi.org/10.1006/ijhc.1996.0091>. Acesso em: 25 set. 2022.

GUARINO, N. Formal ontology in information systems. **Proceedings of the 1st Formal Ontology In Information Systems**, p. 81-97, 1998. Disponível em: https://www.researchgate.net/publication/272169039_Forma_Ontologies_and_Information_Systems. Acesso em: 23 set. 2022.

GUARNIER, T. S. et al. Um Modelo Conceitual Baseado em Ontologia para Doenças Infecciosas com Ênfase em Tuberculose. **Proceedings XIII Seminar on Ontology Research in Brazil (ONTOBRAS 2020)**, Vitória, Brazil, CEUR-WS. 2020. Disponível em: <http://ceur-ws.org/Vol-2728/short5.pdf>. Acesso em: 03 jan. 2023.

GUBERT, C. R. S. et al. Evasão escolar: um fenômeno a ser enfrentado. **Interiorização do ensino superior: protagonismo das universidades estaduais e municipais no desenvolvimento regional**, p. 31, 2018. Disponível em: https://www.udesc.br/arquivos/ceo/id_cpmenu/1887/abruem_2018_15450598880131_1887.pdf#page=31. Acesso em: 30 ago. 2022.

GUIMARÃES, G. C. **Avaliação do PNAES e o direito à permanência no ensino superior nas IFES: O que temos feito?**. Trabalho de Conclusão de Curso. Universidade Federal de

Uberlândia. Graduação em Ciências Sociais. 2020. Disponível em: <https://repositorio.ufu.br/handle/123456789/30584>. Acesso em: 13 abr. 2022.

GUIZZARDI, G. **Desenvolvimento para e com reuso**: Um estudo de caso no domínio de Vídeo sob Demanda. Dissertação de Mestrado. Universidade Federal do Espírito Santo. Mestrado em Informática. 2000. Disponível em: http://www.inf.ufes.br/~gguizzardi/dissertacao_msc.pdf. Acesso em: 01 out. 2022.

GUIZZARDI, G. **Ontological Foundations for Structural Conceptual Model**. CTIT-Centre for Telematics and Information Technology. Doctoral Thesis. University of Twente. 2005. Disponível em: <https://research.utwente.nl/en/publications/ontological-foundations-for-structural-conceptual-models>. Acesso em: 02 abr. 2022.

GUIZZARDI, G. On ontology, ontologies, conceptualizations, modeling languages. **(Meta) Models, Frontiers in Artificial Intelligence and Applications, Databases and Information Systems IV**, p. 18-39, 2007. Disponível em: <https://dl.acm.org/doi/10.5555/1565421.1565425>. Acesso em: 28 set. 2022.

GUIZZARDI, G.; FALBO, R.; GUIZZARDI, R. A importância de Ontologias de Fundamentação para a Engenharia de Ontologias de Domínio: o caso do domínio de Processos de Software. **IEEE Latin America Transactions**, v. 6, n. 3, p. 244-251, 2008. Disponível em: https://nemo.inf.ufes.br/wp-content/papercite-data/pdf/a_importancia_de_ontologias_de_fundamentacao_para_a_engenharia_de_ontologias_de_dominio_o_caso_do_dominio_de_processos_de_software_2008.pdf. Acesso em: 06 nov. 2023.

GUIZZARDI, G. et al. Towards ontological foundations for conceptual modeling: The unified foundational ontology (UFO) story. **Applied ontology**, v. 10, n. 3-4, p. 259-271, 2015. Disponível em: https://nemo.inf.ufes.br/wp-content/papercite-data/pdf/towards_ontological_foundations_for_conceptual_modeling_the_unified_foundational_ontology_ufo_story_2015.pdf. Acesso em: 04 nov. 2023.

GUIZZARDI, G. et al. **Types and taxonomic structures in conceptual modeling**: A novel ontological theory and engineering support. *Data & Knowledge Engineering*, v. 134, p. 101891, 2021. Disponível em: <https://philarchive.org/archive/GUITAT-3>. Acesso em: 22 set. 2023.

GUNEY, Y. Exogenous and endogenous factors influencing students' performance in undergraduate accounting modules. **Accounting Education**, v. 18, n. 1, p. 51-73, 2009. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/09639280701740142>. Acesso em: 20 jul. 2022.

GUO, J.; WANG, Q.; LI, Y. Semi-supervised learning based on convolutional neural network and uncertainty filter for façade defects classification. **Computer-Aided Civil and Infrastructure Engineering**, v. 36, n. 3, p. 302-317, 2021. Disponível em: <https://doi.org/10.1111/mice.12632>. Acesso em: 30 ago. 2022.

GUTIERRE, G. G. **Análise comparativa das taxas de evasão e motivação universitária antes e durante a pandemia de COVID-19**. Trabalho de Conclusão de Curso. Universidade Federal da Paraíba. Curso de Psicologia. 2021. Disponível em: <https://repositorio.ufpb.br/jspui/handle/123456789/21804>. Acesso em: 03 mai. 2022.

HAAV, H. M.; LUBI, T. L. A survey of concept-based information retrieval tools on the web. **Proceedings of the 5th East-European Conference ADBIS**, p. 29-41, 2001. Disponível em:

<https://cdn.cocodoc.com/cocodoc-form-pdf/pdf/289124-fillable-survey-of-concept-based-ir-for-m-mii.pdf>. Acesso em: 24 set. 2022.

HAM, C.; HILL, M. J. **The Policy Process in the Modern Capitalist State**. Harvester Wheatsheaf, 1993.

HARIRI, R. H.; FREDERICKS, E. M.; BOWERS, K. M. Uncertainty in big data analytics: survey, opportunities, and challenges. **Journal of Big Data**, v. 6, n. 1, p. 1-16, 2019. Disponível em: <https://doi.org/10.1186/s40537-019-0206-3>. Acesso em: 23 ago. 2022.

HASAN, M. et al. Current landscape and influence of big data on finance. **Journal of Big Data**, v. 7, n. 1, p. 1-17, 2020. Disponível em: <https://doi.org/10.1186/s40537-020-00291-z>. Acesso em: 20 ago. 2022.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. Unsupervised Learning. The Elements of Statistical Learning. **Springer Series in Statistics**. Springer, New York, NY, 2009. Disponível em: https://doi.org/10.1007/978-0-387-84858-7_14. Acesso em: 31 ago. 2022.

HEINZLE, R. **Um modelo de engenharia do conhecimento para sistemas de apoio à decisão com recursos para raciocínio abduutivo**. Tese de Doutorado. Universidade Federal de Santa Catarina. Doutorado em Engenharia e Gestão do Conhecimento. 2011. Disponível em: <https://repositorio.ufsc.br/handle/123456789/95417>. Acesso em: 20 set. 2022.

HELM, J. M. et al. Machine learning and artificial intelligence: definitions, applications, and future directions. **Current reviews in musculoskeletal medicine**, v. 13, n. 1, p. 69-76, 2020. Disponível em: <https://link.springer.com/article/10.1007/s12178-020-09600-8>. Acesso em: 26 ago. 2022.

HOWLETT, M.; RAMESH, M.; PERL, A. **Studying public policy: Policy cycles and policy subsystems**. Toronto: Oxford university press, 2009.

IFMG. **A evasão escolar no IFMG: diagnóstico e diretrizes da política institucional para a permanência e o êxito dos estudantes**. 2017. Disponível em: <https://www2.ifmg.edu.br/portal/links/relatorio-evasao-completo-rev6.pdf>. Acesso em: 30 abr. 2022.

IFMG. Resolução nº 47 de 17 de dezembro de 2018. Conselho Superior. **Dispõe sobre o Regulamento de Ensino dos Cursos de Graduação do IFMG**. 2018. Disponível em: https://www.ifmg.edu.br/ouropreto/ensino/legislacao-e-normas-de-ensino/regulamentos-de-ensino/Resolucao47_2018RegulamentoEnsinoCursosdeGraduao.pdf. Acesso em: 20 jul. 2022.

IFMG. Resolução nº 9 de 03 de julho de 2020. Conselho Superior. **Dispõe sobre a Aprovação da Política de Assistência Estudantil no âmbito do IFMG**. 2020. Disponível em: <https://www2.ifmg.edu.br/portal/sobre-o-ifmg/conselho-superior/resolucoes/2020-3/resolucao-no-009-2020-aprovacao-da-politica-de-assistencia-estudantil-no-ambito-do-ifmg-e-revogacao-da-resolucao-no-3-2019/view>. Acesso em: 01 jun 2022.

IFMG. **Histórico e Missão**. 2022a. Disponível em: <https://www.ifmg.edu.br/portal/sobre-o-ifmg/historico-e-missao>. Acesso em: 15 jul. 2022.

IFMG. **Diretoria de Assuntos Estudantis**. 2022b. Disponível em: <https://www.ifmg.edu.br/portal/dirae-1/assistencia-estudantil>. Acesso em: 14 jul. 2022.

IFMG. **Relatório de gestão 2021**. 2022c. Disponível em: <https://www.ifmg.edu.br/portal/diretoria-de-desenvolvimento-institucional-ddi/RESOLUON9DE31DEMARODE2022RelatriodeGesto2021.pdf>. Acesso em: 05 jun. 2022.

IMPERATORI, T. K. A trajetória da assistência estudantil na educação superior brasileira. **Serviço Social & Sociedade**, p. 285-303, 2017. Disponível em: <https://repositorio.unb.br/handle/10482/30624>. Acesso em: 10 mai. 2022.

INEP. **Resumo técnico do Censo da Educação Superior 2020**. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2022. Disponível em: https://download.inep.gov.br/publicacoes/institucionais/estatisticas_e_indicadores/resumo_tecnico_censo_da_educacao_superior_2019.pdf. Acesso em: 02 abr. 2021.

IRANZAD, R; LIU, X. A review of random forest-based feature selection methods for data science education and applications. **International Journal of Data Science and Analytics**, p. 1-15, 2024. Disponível em: <https://link.springer.com/article/10.1007/s41060-024-00509-w>. Acesso em: 15 abr. 2025.

JACOB, E. K. Ontologies and the Semantic Web. **Bulletin of the American Society for Information Science and Technology**, v. 29, n. 4, p. 19-19, 2003. Disponível em: <https://doi.org/10.1002/bult.283>. Acesso em: 25 set. 2022.

JOHANN, C. C. **Evasão escolar no Instituto Federal Sul-Rio-Grandense: um estudo de caso no campus Passo Fundo**. Dissertação de Mestrado. Universidade de Passo Fundo. Mestrado em Educação. 2012. Disponível em: <http://tede.upf.br/jspui/handle/tede/739>. Acesso em: 20 abr. 2022.

JORDAN, M. I.; MITCHELL, T. Machine learning: Trends, perspectives, and prospects. **Science**, v. 349, n. 6245, p. 255-260, 2015. Disponível em: <https://www.science.org/doi/abs/10.1126/science.aaa8415>. Acesso em: 29 ago. 2022.

JORGE, E. M. F. J. et al. Arquitetura da informação analítica para integração de dados da pesquisa e pós-graduação: um estudo de caso da Universidade do Estado da Bahia. **Informação & Informação**, v. 25, n. 1, p. 115-140, 2020. Disponível em: <https://www.uel.br/revistas/uel/index.php/informacao/article/view/36009>. Acesso em: 02 out. 2022.

JOSE, A. R.; BROILO, C. L.; ANDREOLI, G. S. **A evasão na Unipampa – diagnosticando processos, acompanhando trajetórias e itinerários de formação**. Universidade Federal do Pampa, 2011. Disponível em: https://sites.unipampa.edu.br/formacao/files/2010/07/relatorio_final_evasao-na-unipampa_ou_t20111.pdf. Acesso em: 12 out. 2022.

JUCA, S. C. S. et al. Acesso, permanência e êxito no Ensino Superior: análise do desempenho acadêmico e da evasão de estudantes no IFCE. **Revista Thema**, v. 16, n. 1, p. 115-128, 2019. Disponível em: <https://periodicos.ifsul.edu.br/index.php/thema/article/view/1170>. Acesso em: 19 jul. 2022.

KEET, C. M. **An introduction to ontology engineering**. Cape Town, South Africa: Maria Keet, 2020. Disponível em: <https://people.cs.uct.ac.za/~mkeet/files/OEbook.pdf>. Acesso em: 29 set. 2022.

KHAMDAMOV, U. et al. Models of integration of higher education management information systems. **2021 International Conference on Information Science and Communications Technologies (ICISCT)**, IEEE, 2021. Disponível em: <https://ieeexplore.ieee.org/abstract/document/9670171>. Acesso em: 16 out. 2022.

KHANDKER, S. R.; KOOLWAL, G. B.; SAMAD, H. A. Handbook on impact evaluation: quantitative methods and practices. **Washington: The World Bank**, 2010. Disponível em: <https://documents1.worldbank.org/curated/en/650951468335456749/pdf/520990PUB0EPI1101Official0Use0Only1.pdf>. Acesso em: 02 mai. 2024.

KOTIRANTA, P.; JUNKKARI, M.; NUMMENMAA, J. Performance of Graph and Relational Databases in Complex Queries. **Applied Sciences**, v. 12, n. 13, 2022. Disponível em: <https://www.mdpi.com/2076-3417/12/13/6490>. Acesso em: 24 abr. 2025.

KOWALSKI, A. V. **Os (des)caminhos da política de assistência estudantil e o desafio da garantia de direito**. Tese de Doutorado. PUCRS. Doutorado em Serviço Social. 2012. Disponível em: <https://tede2.pucrs.br/tede2/handle/tede/521>. Acesso em: 30 jun. 2022.

LIES, J. Marketing intelligence and big data: Digital marketing techniques on their way to becoming social engineering techniques in marketing. **IJIMAI**, v. 5, n. 5, 2019. Disponível em: <https://reunir.unir.net/handle/123456789/12532>. Acesso em: 24 ago. 2022.

LIMA, Q.; PAZINATO, M. S. Fatores que interferem no desempenho acadêmico em cursos de licenciatura em ciência da natureza. **Anais do Salão Internacional de Ensino, Pesquisa e Extensão**, v. 10, n. 2, 2018. Disponível em: https://guri.unipampa.edu.br/uploads/evt/arg_trabalhos/18121/seer_18121.pdf. Acesso em: 30 jul. 2022.

LIMA, W. A. S. **Permanecer no Ensino Superior, eis a questão: estratégias de implementação efetiva de políticas públicas de Assistência Estudantil**. Dissertação de Mestrado. Universidade Federal da Bahia. Mestrado Profissional em Administração. 2017. Disponível em: <https://repositorio.ufba.br/handle/ri/24271>. Acesso em: 11 jul. 2022.

LIMA, W. A. S.; FERREIRA, L. C. Mapeamento e avaliação das políticas públicas de assistência estudantil nas universidades federais brasileiras. **Revista Meta: Avaliação, Brasil**, v. 8, n. 22, p. 116-148, 2016. Disponível em: https://web.archive.org/web/20170922023506id_/http://revistas.cesgranrio.org.br/index.php/metaavaliacao/article/viewFile/870/pdf1. Acesso em: 12 jul. 2022.

LINKE, E. C.; NOGUEIRA, B. C.; LINKE, E. C. A evasão escolar no ensino técnico profissionalizantes. **Anais do 22º Seminário Interinstitucional de Ensino, Pesquisa e Extensão**, p. 01-14, 2017. Disponível em: https://www.unicruz.edu.br/seminario/anais/anais-2017/XXII%20SEMIN%C3%81RIO%20INTERINSTITUCIONAL%202017%20-%20ANAIIS/GRADUA%C3%87%C3%83O%20-%20TRABALHOS%20COMPLETOS_Ci%C3%A4ncias%20Sociais%20e%20Humanidades/A%20Evas%C3%A3o%20escolar%20a%20n%C3%ADvel%20t%C3%A9cnico%20profissionalizante.pdf. Acesso em: 19 ago. 2022.

LOPES, A. E. M. P. Limitadores ao progresso de um ecossistema empreendedor à luz da Gestão da Tecnologia da Informação. **Revista Interdisciplinar Científica Aplicada**, v. 14, n. 4, p. 39-61, 2020. Disponível em: <https://portaldeperiodicos.animaeducacao.com.br/index.php/rica/article/view/18025>. Acesso em: 25 out. 2022.

LOPES, B.; AMARAL, J. N. **Políticas Públicas: conceitos e práticas**. Belo Horizonte: Sebrae, 2008.

LOPES JÚNIOR, A. et al. Using terms and informal definitions to classify domain entities into top-level ontology concepts: An approach based on language models. *Knowledge-Based Systems*, v. 265, p. 110385, 2023. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0950705123001351>. Acesso em: 20 abr. 2025.

LOURENÇO, A. V. M. et al. O fenômeno da evasão no ensino superior: um estudo de caso nos cursos de Administração da UNIGRANRIO. **XIII Colóquio de Gestão Universitária em Américas**. 2013. Disponível em: <https://repositorio.ufsc.br/handle/123456789/113103>. Acesso em: 24 mar. 2022.

MACEDO, S. M. S.; BARBOSA, R. R. Gestão da informação, da tecnologia da informação de comportamentos e valores relativos à informação em instituições de ensino superior (IES) de Belo Horizonte. **Brazilian Journal of Information Science**, v. 7, n. 1, p. 137-153, 2013. Disponível em: <https://dialnet.unirioja.es/servlet/articulo?codigo=4331896>. Acesso em: 23 out. 2022.

MACHADO, G. C.; OLIVEIRA, C. A.; FREITAS, T. A. Avaliação do impacto dos benefícios de assistência estudantil sobre o desempenho acadêmico: o caso da Universidade Federal do Rio Grande. **Planejamento e Políticas Públicas**, n. 55, 2020. Disponível em: <https://www.ipea.gov.br/ppp/index.php/PPP/article/view/1091>. Acesso em: 30 out. 2022.

MACIEL, D. A. S. **Políticas Públicas e Democratização Educacional: Acesso e Permanência no Ensino Superior através do Programa Reuni**. Dissertação de Mestrado. Universidade Federal da Bahia. Mestrado em Estudos Interdisciplinares sobre a Universidade. 2014. Disponível em: <https://repositorio.ufba.br/handle/ri/16465>. Acesso em: 20 mai. 2022.

MADDALENA, L.; BAIÃO, F. OntoCovid: Aplicando SABIÓ para a modelagem. **CEUR Workshop Proceedings**, 2021. Disponível em: <http://ceur-ws.org/Vol-3050/Short5.pdf>. Acesso em: 28 out. 2022.

MAGRANI, E. **A internet das coisas**. Rio de Janeiro: FGV, 2018.

MALIN, A. M. B. **Estado e sociedade da informação no Brasil: uma investigação sobre o papel da informação na gestão pública**. Tese de Doutorado. Universidade Federal do Rio de Janeiro. Doutorado em Ciência da Informação. 2003. Disponível em: <http://repositorio.ibict.br/handle/123456789/677>. Acesso em: 25 out. 2022.

MAMEDE, S. et al. Determinantes Psicológicos do Desempenho Acadêmico em Ciências Contábeis: Evidências do Brasil. **Brazilian Business Review**, edição especial, p. 54-75, 2015. Disponível em: <https://core.ac.uk/download/pdf/234099928.pdf>. Acesso em: 23 nov. 2022.

MARCACINI, R. M. **Aprendizado de máquina com informação privilegiada: abordagens para agrupamento hierárquico de textos**. Tese de Doutorado. Universidade de São Paulo. Doutorado em Ciência da Computação e Modelagem Matemática. 2014. Disponível em: <https://www.teses.usp.br/teses/disponiveis/55/55134/tde-05082015-094733/en.php>. Acesso em: 03 set. 2022.

MARCHAND, D. A.; KETTINGER, W. J.; ROLLINS, J. **Information orientation: the link to business performance**. Nova York: University Oxford, 2000.

MARCONI, M. A.; LAKATOS, E. M. **Metodologia do trabalho científico**: projetos de pesquisa/pesquisa bibliográfica/teses de doutorado, dissertações de mestrado, trabalhos de conclusão de curso. São Paulo: Atlas, 2017.

MARTINS, C. B. N. **Evasão de alunos nos cursos de graduação em uma instituição de ensino superior**. Dissertação de Mestrado. Fundação Dr. Pedro Leopoldo. Mestrado em Administração. 2007. Disponível em: https://fpl.edu.br/2018/media/pdfs/mestrado/dissertacoes_2007/dissertacao_cleidis_beatriz_nogueira_martins_2007.pdf. Acesso em: 11 jul. 2022.

MARTINS, F. N. C. **Políticas de assistência estudantil universitária**: definições, limites e possibilidades na perspectiva de um atendimento de qualidade. Dissertação de Mestrado. Universidade Lusófona de Humanidades e Tecnologias. Mestrado em Ciências da Educação. 2022. Disponível em: <https://recil.ensinulusofona.pt/handle/10437/12915>. Acesso em: 7 jul. 2022.

MARTINS, P. F. M.; SILVA, E. G.; MAURICIO, N. M. M. A história da Assistência Estudantil no Ensino Superior Brasileiro: programa nacional de assistência estudantil e o aumento das classes "D" e "E" nas universidades federais. **Revista Observatório**, v. 5, n. 6, p. 886-911, 2019. Disponível em: <https://sistemas.uft.edu.br/periodicos/index.php/observatorio/article/view/7493>. Acesso em: 3 jul. 2022.

MARTINS, S. C. **Gestão da Informação**: estudo comparativo de modelos sob a ótica integrativa dos recursos de informação. Dissertação de Mestrado. Universidade Federal Fluminense. Mestrado em Ciência da Informação. 2014. Disponível em: <https://app.uff.br/riuff/handle/1/10725>. Acesso em: 18 out. 2022.

MARTINS, Z. B.; MARINHO, S. V. Relação das variáveis concernentes ao desempenho acadêmico: um estudo com alunos de graduação em ciências contábeis. **Revista Universo Contábil**, v. 15, n. 1, p. 27-48, 2019. Disponível em: <https://bu.furb.br/ojs/index.php/universocontabil/article/view/6604>. Acesso em: 02 ago. 2022.

MATTA, K. W. **Evasão universitária estudantil**: precursores psicológicos do trancamento de matrícula por motivo de saúde mental. Dissertação de Mestrado. Universidade de Brasília. Mestrado em Psicologia Clínica e Cultura. 2011. Disponível em: <https://repositorio.unb.br/handle/10482/9338>. Acesso em: 28 jul. 2022.

MCGEE, J.; PRUSAK, L. **Gerenciamento estratégico da informação**. Rio de Janeiro: Campus, 1994.

MEC. Ministério da Educação. **Instituições da rede federal**. 2024. Brasília, DF: MEC. Disponível em: <https://www.gov.br/mec/pt-br/areas-de-atuacao/ept/rede-federal>. Acesso em: 22 nov. 2024.

MELO, E. C. P. et al. Reuni como política de expansão das universidades federais e os cortes do Governo Federal. **Research, Society and Development**, v. 9, n. 12, p. e39491211400, 2020. Disponível em: <https://rsdjournal.org/index.php/rsd/article/view/11400>. Acesso em: 9 jul. 2022.

MELO, J. R. T. **O Programa Nacional de Assistência Estudantil-PNAES e os desafios para a inclusão social face à política da dor**. Tese de Doutorado. Universidade Federal Rural da Amazônia. Mestrado em Educação. 2019. Disponível em: <http://www.repositorio.ufra.edu.br:8080/jspui/handle/123456789/1181>. Acesso em: 10 jul. 2022.

MELLO, S. P. T.; SANTOS, E. G. Diagnóstico e alternativas de contenção da evasão no curso de Administração em uma universidade pública no Sul do Brasil. **Revista Gestão Universitária na América Latina-GUAL**, v. 5, n. 3, p. 67-80, 2012. Disponível em: <https://www.redalyc.org/pdf/3193/319327516005.pdf>. Acesso em: 27 ago. 2022.

MENDONÇA, F. M.; ZAIDAN, F. H. Ontologias para organização da informação em processos de transformação digital. **Em Questão**, v. 25, n. 1, p. 295-320, 2019. Disponível em: <https://www.seer.ufrgs.br/index.php/EmQuestao/article/view/80311/50055>. Acesso em: 21 mar. 2023.

MIRAFETABZADEH, S. M. et al. K-means and alternative clustering methods in modern power systems. **IEEE Access**, v. 11, p. 119596-119633, 2023. Disponível em: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&number=10295469>. Acesso em: 15 abr. 2025.

MIRANDA, G. J. et al. Determinantes do desempenho acadêmico na área de negócios. **Revista Meta: Avaliação**, v. 7, n. 20, p. 175-209, 2015. Disponível em: https://www.researchgate.net/profile/Gilberto-Miranda/publication/283719584_Determinantes_do_Desempenho_Academico_na_Area_de_Negocios/links/5644a93708ae54697fb8040d/Determinantes-do-Desempenho-Academico-na-Area-de-Negocios.pdf. Acesso em: 20 jul. 2022.

MIRANDA, G. J.; NOVA, S. P. C. C.; CORNACCHIONE JÚNIOR, E. B. Ao mestre com carinho: relações entre as qualificações docentes e o desempenho discente em Contabilidade. **Revista Brasileira de Gestão de Negócios**, v. 15, p. 462-480, 2013. Disponível em: <https://www.scielo.br/rbgn/a/JVfDFmbH3RZWw3HJ7Qf4mLN/abstract/?lang=pt>. Acesso em: 05 ago. 2022.

MITCHELL, T. **Machine Learning**. McGraw-Hill, 1997.

MOMM, A. M. P.; MOMM, S. F. **A evasão escolar no curso superior de tecnologia em Jaraguá do Sul**. Trabalho de Conclusão de Curso. Instituto Federal de Santa Catarina. e Especialização em Gestão Pública na Educação Profissional. 2020. Disponível em: <https://repositorio.ifsc.edu.br/handle/123456789/1417>. Acesso em: 17 ago. 2022.

MONARD, M. C.; BARANAUSKAS, J. A. Conceitos sobre aprendizado de máquina. **Sistemas inteligentes - Fundamentos e aplicações**, v. 1, n. 1, 2003. Disponível em: <https://dcm.ffclrp.usp.br/~augusto/publications/2003-sistemas-inteligentes-cap4.pdf>. Acesso em: 26 ago. 2022.

MOON, K.; BLACKMAN, D. A guide to understanding Social Science Research for Natural Scientists. **Conservation Biology**, v. 28, n. 5, 2014. Disponível em: <https://conbio.onlinelibrary.wiley.com/doi/abs/10.1111/cobi.12326>. Acesso em: 18 jul. 2024.

MOREIRA, L. J.; GOMES, T. S. **Determinantes do desempenho acadêmico do estudante trabalhador e do trabalhador estudante**. Trabalho de Conclusão de Curso. Universidade Tecnológica Federal do Paraná. Curso de Ciências Contábeis. 2018. Disponível em: http://repositorio.roca.utfpr.edu.br/jspui/bitstream/1/14226/1/PB_COCTB_2018_2_13.pdf. Acesso em: 01 ago. 2022.

MUNHOZ, A. M. H. **Uma análise multidimensional da relação entre inteligência e desempenho acadêmico em universitários ingressantes**. Tese de Doutorado. Universidade Estadual de Campinas. Doutorado em Educação. 2004. Disponível em:

https://www.psiem.fe.unicamp.br/pf-psiem/hernandezmunhoz_aliciamaria_d.pdf. Acesso em: 19 jul. 2022.

MUSSLINER, B. O. et al. O problema da evasão universitária: um desafio à democratização do ensino superior público. **Brazilian Journal of Development**, v. 7, n. 4, p. 42738-42758, 2021. Disponível em: <https://brazilianjournals.com/index.php/BRJD/article/view/28961>. Acesso em: 10 mai. 2022.

NAGAI, N. P.; CARDOSO, A. L. J. A evasão universitária: Uma análise além dos números. **Revista Estudo & Debate**, v. 24, n. 1, 2017. Disponível em: <http://www.univates.com.br/revistas/index.php/estudoedebate/article/view/1271>. Acesso em: 28 jul. 2022.

NAPOLEON, D.; PAVALAKODI, S. A new method for dimensionality reduction using k-means clustering algorithm for high dimensional data set. **International Journal of Computer Applications**, v. 13, n. 7, p. 41-46, 2011. Disponível em: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=f55569a0a7484b6086185f8c4119b8246e5da4da>. Acesso em: 06 jan. 2023.

NAQA, I.; MURPHY, M. J. What is machine learning?. **Machine learning in radiation oncology**. Springer, Cham, p. 3-11, 2015. Disponível em: https://link.springer.com/chapter/10.1007/978-3-319-18305-3_1. Acesso em: 26 ago. 2022.

NARDINI, E. F. et al. Política de estímulo à iniciação científica: impacto no coeficiente de rendimento de graduandos em Odontologia. **Revista da ABENO**, v. 19, n. 1, p. 33-39, 2019. Disponível em: <https://revabeno.emnuvens.com.br/revabeno/article/view/619>. Acesso em: 05 ago. 2022.

NASCIMENTO, L. M. **Utilizando linked data para publicação e cruzamento de dados governamentais abertos**. Trabalho de Conclusão de Curso. Universidade Federal Fluminense. Graduação em Ciência da Computação. 2018. Disponível em: https://oasisbr.ibict.br/vufind/Record/UFF-2_af4b82e468318145cb68505765930f21. Acesso em: 12 jan. 2023.

NOY, N. F.; MCGUINNESS, D. L. Ontology development 101: A guide to creating your first ontology. **Stanford Knowledge Systems Laboratory Technical Report KSL01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880**. 2001. Disponível em: https://corais.org/sites/default/files/ontology_development_101_aguide_to_creating_your_first_ontology.pdf. Acesso em: 17 out. 2022.

NIQUINI, R. P. et al. Características do trabalho de estudantes universitários associadas ao seu desempenho acadêmico. **Educação em Revista**, v. 31, p. 359-381, 2015. Disponível em: <https://www.scielo.br/ij/edur/a/VgbNZW56cYgfQdbDzKw7GVL/abstract/?lang=pt>. Acesso em: 05 ago. 2022.

NOGUEIRA, D. R. Desempenho acadêmico x estilos de aprendizagem segundo Honey-Alonso: uma análise com alunos do curso de Ciências Contábeis. **Revista Espaço Acadêmico**, v. 12, n. 137, p. 80-89, 2012. Disponível em: <https://periodicos.uem.br/ojs/index.php/EspacoAcademico/article/view/16719>. Acesso em: 18 jul. 2022.

NOGUEIRA, W. B. **Avaliação de desempenho de uma usina solar fotovoltaica com base no agrupamento k-means**. Trabalho de Conclusão de Curso. Universidade Federal do Ceará. Bacharelado em Engenharia Elétrica. 2022. Disponível em: <https://repositorio.ufc.br/handle/riufc/65886>. Acesso em: 03 set. 2022.

NONAKA, I.; TAKEUCHI, H. **Gestão do Conhecimento**. Porto Alegre: Bookman, 2008.

NONATO, B. F. et al. Mudanças no perfil dos estudantes da UFMG: desafios para a prática docente. **Revista Docência do Ensino Superior**, v. 10, p. 1–21, 2020. Disponível em: <https://periodicos.ufmg.br/index.php/rdes/article/view/20463>. Acesso em: 1 abr. 2022.

NONATO, R. S.; AGANETTE, E. C. Gestão da informação: rumo a uma proposta de definição atual e consensual para o termo. **Perspectivas em Ciência da Informação**, v. 27, p. 133-159, 2022. Disponível em: <https://www.scielo.br/j/pci/a/X3WNQHDXTFSNXNspLV73DRw/abstract/?lang=pt>. Acesso em: 24 out. 2022.

NORMAN, A. T. **Aprendizado de Máquina em Ação**. Rússia: ЛитРес, 2021.

OLIVEIRA, A. J. et al. Programa Reuni nas instituições de Ensino Superior federal [IFES] brasileiras: um estudo da eficiência operacional por meio da análise envoltória de dados [DEA] no período de 2006 a 2012. **Race: revista de administração, contabilidade e economia**, v. 13, n. 3, p. 1179-1210, 2014. Disponível em: <https://dialnet.unirioja.es/servlet/articulo?codigo=5160886>. Acesso em: 29 jun. 2022.

OLIVEIRA, E. S. G.; VASCONCELLOS, P. R. **Grupo focal em informática na educação: diálogo, conflito, consenso?** In: PIMENTEL, M.; SANTOS, E. (Org.) Metodologia de pesquisa científica em Informática na Educação: abordagem qualitativa. Porto Alegre: SBC, 2021.

OLIVEIRA, M. A. M.; DEUS, L. A.; CAMPOS, M. S. Avaliação do PNAES sobre a permanência de estudantes beneficiados em um campus universitário público federal. **Revista Eletrônica de Educação**, v. 18, n. 1, p. e6043126-e6043126, 2024. Disponível em: <https://reveduc.ufscar.br/index.php/reveduc/article/download/6043/1677>. Acesso em: 17 abr. 2025.

OLIVEIRA, M. S. **Metodologia de seleção de features não supervisionada para clustering em conjunto de dados de alta dimensionalidade**. Dissertação de Mestrado. Universidade Federal de Pernambuco. Mestrado em Ciência da Computação. 2018. Disponível em: <https://repositorio.ufpe.br/handle/123456789/33642>. Acesso em: 02 set. 2022.

OLIVEIRA, P. C. et al. Gestão da informação ou gestão do conhecimento? Percepção de profissionais da indústria de software de Santa Catarina. **Perspectivas em Gestão & Conhecimento**, v. 4, n. 2, p. 67-89, 2014. Disponível em: <https://dialnet.unirioja.es/servlet/articulo?codigo=4929754>. Acesso em: 18 out. 2022.

OLIVEIRA, P. S. **Relação entre orçamento e questionamentos no portal da transparência: análise de dados das Universidades federais do Brasil entre 2014 e 2018**. Trabalho de Conclusão de Curso. Universidade de Brasília. Graduação em Ciências Contábeis. 2019. Disponível em: <https://bdm.unb.br/handle/10483/27535>. Acesso em: 12 set. 2022.

ORESKI, D.; PIHIR, I.; KONECKI, M. CRISP-DM process model in educational setting. **Economic and Social Development: Book of Proceedings**, p. 19-28, 2017. Disponível em: https://www.bib.irb.hr/983301/download/983301.Book_of_Proceedings_esdPrague_2017.pdf#page=28. Acesso em: 24 ago. 2022.

ORLANDO, J. P. **ONTO-Analyst: um método extensível para a identificação e visualização de anomalias em ontologias**. Tese de Doutorado. Universidade de São Paulo. Doutorado em

Ciência da Computação e Matemática Computacional. 2017. Disponível em: <https://www.teses.usp.br/teses/disponiveis/55/55134/tde-07122017-102344/en.php>. Acesso em: 17 out. 2022.

PAIVA, M. O. A.; LOURENÇO, A. A. Rendimento acadêmico: Influência do autoconceito e do ambiente de sala de aula. **Psicologia: Teoria e Pesquisa**, v. 27, p. 393-402, 2011. Disponível em: <https://www.scielo.br/j/ptp/a/V7GrLdg7cv4XYkN7FJZYTwd/?lang=pt&format=html>. Acesso em: 02 ago. 2022.

PAULA, C. H.; ALMEIDA, F. M. O programa Reuni e o desempenho das Ifes brasileiras. **Ensaio: Avaliação e Políticas Públicas em Educação**, v. 28, p. 1054-1075, 2020. Disponível em: <https://www.scielo.br/j/ensaio/a/5pvgF4sGMQsn89ZYSYfWHsh/abstract/?lang=pt>. Acesso em: 8 jul. 2022.

PAULA, M. F. C. Políticas de democratização da educação superior brasileira: limites e desafios para a próxima década. **Revista da Avaliação da Educação Superior (Campinas)**, v. 22, p. 301-315, 2017. Disponível em: <https://www.scielo.br/j/aval/a/KYs6H9L5YpppTCZHpHGd8SK/abstract/?lang=pt>. Acesso em: 29 mai. 2022.

PAVÃO, J. A.; FERREIRA, J. L. N.; FERREIRA, M. M. A Influência da Qualidade de Vida e dos Fatores Internos da Instituição no Desempenho Acadêmico. **Contabilometria**, v. 6, n. 1, 2019. Disponível em: <https://revistas.fucamp.edu.br/index.php/contabilometria/article/view/1300>. Acesso em: 23 ago. 2022.

PEREIRA, D. B. O processo de monitoramento de políticas públicas de saúde: o caso do Núcleo Estratégia Saúde da Família/RN. **IV Encontro Brasileiro de Administração Pública**, p. 649-664, 2017. Disponível em: https://www.academia.edu/35083508/O_processo_de_monitoramento_de_pol%C3%ADticas_p%C3%BAblicas_de_sa%C3%BAde_o_caso_do_N%C3%BAcleo_Estrat%C3%A9gia_Sa%C3%BAde_da_Fam%C3%ADlia_RN. Acesso em: 10 abr. 2024.

PEREIRA, F. C. M. **Comportamento informacional na tomada de decisão**: proposta de Modelo Integrativo. Tese de Doutorado. Universidade Federal de Minas Gerais. Doutorado em Ciência da Informação. 2011. Disponível em: <https://repositorio.ufmg.br/handle/1843/BUOS-8PGLKJ>. Acesso em: 29 out. 2022.

PEREIRA, L. T.; ALMEIDA, M. S. A definição do perfil de vulnerabilidade para acesso à política de permanência no ensino superior. **Revista GUAL**, v. 8, n. 1, p. 132-154, 2015. Disponível em: <https://www.redalyc.org/pdf/3193/319338454008.pdf>. Acesso em: 23 mai. 2022.

PEREIRA, P. A. P.; STEIN, R. H. Política social: universalidade versus focalização. Um olhar sobre a América Latina. In: BEHRING, Elaine Rossetti et al. (Orgs.). **Capitalismo em crise, política social e direitos**. São Paulo: Cortez, 2010. p. 106-30. Disponível em: <http://ria.ufrn.br:8080/jspui/handle/123456789/1112>. Acesso em: 30 mai. 2022.

PESSANHA, C. P.; BAX, M. P. Implementando o prontuário eletrônico OpenEHR em Sistemas Gestores de Conteúdo: similitude entre arquétipos e conteúdos. **XVII Encontro Nacional de Pesquisa em Ciência da Informação**, 2016. Disponível em: https://www.brapci.inf.br/repositorio/2016/11/pdf_4aed3cd220_0000021715.pdf. Acesso em: 21 mar. 2023.

PETTIGREW, K. E.; FIDEL, R.; BRUCE, H. Conceptual frameworks in information behavior. **Annual review of information science and technology (ARIST)**, v. 35, n. 43-78, 2001. Disponível em: <http://faculty.washington.edu/fidelr/RayaPubs/ConceptualFrameworks.pdf>. Acesso em: 27 out. 2022.

PIETSCH, W. **On the Epistemology of Data Science: Conceptual Tools for a New Inductivism**. Springer Nature, 2021.

PINHEIRO, E. B. A trajetória da assistência estudantil no contexto da Política de Educação: educação superior, assistência estudantil e o PNAES. **Anais do XVI Encontro Nacional de Pesquisadores em Serviço Social**, v. 16, n. 1, 2018. Disponível em: <https://periodicos.ufes.br/abepss/article/view/22948>. Acesso em: 11 set. 2022.

PINHEIRO, L. V. R. Itinerários epistemológicos da instituição e constituição da Informação em Arte no campo interdisciplinar da Museologia e da Ciência da Informação. **Revista Museologia e Patrimônio**, v.1, n.1, p. 9-17, 2008. Disponível em: <https://ridi.ibict.br/handle/123456789/9>. Acesso em: 29 dez. 2022.

PINHEIRO, M. M. S. **Políticas públicas baseadas em evidências (PPBEs): delimitando o problema conceitual**. Texto para Discussão, Instituto de Pesquisa Econômica Aplicada, 2020. Disponível em: https://repositorio.ipea.gov.br/bitstream/11058/9915/1/td_2554.pdf. Acesso em: 10 jun. 2023.

PINTO, D.; MOLINA, L. G.; PALETTA, F. C. Uso das tecnologias da informação e comunicação na gestão da informação e do conhecimento nas organizações. **Perspectivas em Gestão & Conhecimento**, v. 12, n. 1, p. 80-96, 2022. Disponível em: <https://www.eca.usp.br/acervo/producao-academica/003074162.pdf>. Acesso em: 24 out. 2022.

PUTRAMA, I. M.; MARTINEK, P. Heterogeneous data integration: Challenges and opportunities. **Data in Brief**, v. 56, n. 110853, p. 1-23. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2352340924008175?via%3Dihub>. Acesso em: 15. abr. 2025.

QUEIROZ, E. V. **Análise da qualidade do sono de estudantes universitários da área de saúde e sua relação com desempenho acadêmico**. Trabalho de Conclusão de Curso. Centro Universitário de Brasília. Curso de Medicina. 2021. Disponível em: <https://repositorio.uniceub.br/jspui/handle/prefix/15133>. Acesso em: 30 jul. 2022.

QUEIROZ, Z. C. L. S. et al. A lei de cotas na perspectiva do desempenho acadêmico na Universidade Federal de Uberlândia. **Revista Brasileira de Estudos Pedagógicos**, v. 96, p. 299-320, 2015. Disponível em: <https://www.scielo.br/j/rbeped/a/wTypQhvyvDpm7sz6MJDRQTD/?lang=pt&format=html>. Acesso em: 19 jul. 2022.

RABELO, D. R.; GOES, B. T. A era da Fisioterapia Baseada em Evidências: o despertar de uma nova forma de ser fisioterapeuta. **Journal of Evidence-Based Healthcare**, v. 1, n. 1, p. 34-39, 2019. Disponível em: <https://www5.bahiana.edu.br/index.php/evidence/article/view/2205/2843>. Acesso em: 13 jan. 2023.

RAMOS, J. L. C. et al. CRISP-EDM: uma proposta de adaptação do Modelo CRISP-DM para mineração de dados educacionais. **Anais do XXXI Simpósio Brasileiro de Informática na Educação**. SBC, p. 1092-1101, 2020. Disponível em: <https://sol.sbc.org.br/index.php/sbie/article/view/12865>. Acesso em: 24 ago. 2022.

RAMOS, M. Aspectos conceituais e metodológicos da avaliação de políticas e programas sociais. **Revista Planejamento e Políticas Públicas**, Brasília, DF, n. 32, p. 95-114, jan./jun. 2009. Disponível em: <http://www.ipea.gov.br/ppp/index.php/PPP/article/view/11>. Acesso em: 9 abr. 2024.

RANGEL, J. R.; MIRANDA, G. J. Desempenho acadêmico e o uso de redes sociais. **Sociedade, Contabilidade e Gestão. Minas Gerais**, v. 11, n. 2, p. 139-154, 2016. Disponível em: <http://atena.org.br/revista/ojs-2.2.3-08/index.php/ufrj/article/viewArticle/2786>. Acesso em: 01 ago. 2022.

RAPHAELL, B. **Desmistificando termos em Machine Learning - tipos de aprendizado**. Alura, 2021. Disponível em: <https://www.alura.com.br/artigos/desmistificando-termos-machine-learning-tipos-aprendizado>. Acesso em: 15 ago. 2022.

RECKTENVALD, M.; MATTEI, L.; PEREIRA, V. A. Avaliando o Programa Nacional de Assistência Estudantil (PNAES) sob a ótica das epistemologias. **Revista da Avaliação da Educação Superior (Campinas)**, v. 23, p. 405-423, 2018. Disponível em: <https://www.scielo.br/j/aval/a/8rJtwcBryJG67DhKZ7grXFw/abstract/?lang=pt>. Acesso em: 22 nov. 2022.

RECTOR, A. et al. OWL pizzas: Practical experience of teaching OWL-DL: Common errors & common patterns. **International Conference on Knowledge Engineering and Knowledge Management**. Springer, Berlin, Heidelberg, p. 63-81, 2004. Disponível em: https://link.springer.com/chapter/10.1007/978-3-540-30202-5_5. Acesso em: 17 out. 2022.

REIS, E. M.; GONÇALVES, A. O. Programa de promoção à permanência: perfil dos discentes do Instituto Federal de Ciência e Tecnologia de Brasília. **HOLOS**, v. 3, p. 379-394, 2021. Disponível em: <https://www2.ifrn.edu.br/ojs/index.php/HOLOS/article/view/4044>. Acesso em: 12 fev. 2022.

REIS, P. R. C. et al. Impactos das aposentadorias e pensões no nível de bem-estar social dos domicílios de Minas Gerais. **Revista Contabilidade e Finanças**, São Paulo, v. 26, n. 67, p. 106-118, jan./abr. 2015. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1519-70772015000100106&lng=en&nrm=iso. Acesso em: 26 mar. 2024.

RESENDE, A. C. C.; OLIVEIRA, A. M. H. C. Avaliando resultados de um programa de transferência de renda: o impacto do Bolsa-Escola sobre os gastos das famílias brasileiras. **Estudos Econômicos (São Paulo)**, v. 38, p. 235-265, 2008. Disponível em: <https://www.scielo.br/j/ee/a/QhRzNk7V9gwrYPxpBMxP6Rc/>. Acesso em: 10 abr. 2024.

RIBEIRO, E. M. B.; PEIXOTO, A. L. A.; BASTOS, A. V. B.. Interação entre estudantes cotistas e não cotistas e sua influência na integração social e desempenho acadêmico na universidade. **Estudos de Psicologia**, v. 22, n. 4, p. 401-411, 2017. Disponível em: <https://www.redalyc.org/pdf/261/26157045007.pdf>. Acesso em: 20 jul. 2022.

RIBEIRO, L. S. et al. **Uma abordagem semântica para seleção de atributos no processo de KDD**. Dissertação de Mestrado. Universidade Federal da Paraíba. Mestrado em Informática. 2010. Disponível em: <https://repositorio.ufpb.br/jspui/handle/tede/6048>. Acesso em: 01 jul. 2022.

RISTOFF, D. O novo perfil do campus brasileiro: uma análise do perfil socioeconômico do estudante de graduação. **Avaliação**, v. 19, n. 3. 2014. Disponível em:

<https://www.scielo.br/j/aval/a/yQz6tVyGStDkzSMZcVpkTbT/?lang=pt>. Acesso em: 28 dez. 2022.

RODRIGUES, B. C. O. et al. Determinantes do desempenho acadêmico dos alunos dos cursos de Ciências Contábeis no ensino à distância. **Enfoque: Reflexão Contábil**, v. 35, n. 2, p. 139-153, 2016. Disponível em: <https://www.redalyc.org/pdf/3071/307146810010.pdf>. Acesso em: 01 ago. 2022.

RODZI, N. A. H. M.; OTHMAN, M. S.; YUSUF, L. M. Significance of data integration and ETL in business intelligence framework for higher education. **2015 International Conference on Science in Information Technology (ICSITech)**, IEEE, 181-186, 2015. Disponível em: <https://ieeexplore.ieee.org/abstract/document/7407800/>. Acesso em: 19 fev. 2023.

ROSENBAUM, P. R.; RUBIN, D. B. The central role of the propensity score in observational studies for causal effects. **Biometrika**, v. 70, n. 1, p. 41-55, 1983. Disponível em: <https://academic.oup.com/biomet/article/70/1/41/240879?login=false>. Acesso em: 29 mar. 2024.

ROZADOS, H. B. F. A ciência da informação em sua aproximação com as ciências cognitivas. **Em Questão**, v. 9, n. 1, p. 79-94, 2003. Disponível em: <https://www.seer.ufrgs.br/EmQuestao/article/view/62>. Acesso em: 27 out. 2022.

RUA, M. G. **Políticas Públicas**. Florianópolis: Departamento de Ciências da Administração. Brasília/DF, CAPES, UAB, 2009. Disponível em: https://cesad.ufs.br/ORBI/public/uploadCatalogo/10551017022012Políticas_Publicas_Aula_1.pdf. Acesso em: 28 mar. 2024.

SACCARO, A.; FRANÇA, M. T. A.; JACINTO, P. A. Retensão e evasão no ensino superior brasileiro: uma análise dos efeitos da bolsa permanência do PNAES. **44º Encontro Nacional de Economia-Anpec**, 2016, Brasil., 2016. Disponível em: https://repositorio.pucrs.br/dspace/bitstream/10923/10648/2/Retensao_e_evasao_no_ensino_superior_brasileiro_uma_analise_dos_efeitos_da_bolsa_permanencia_do_PNAES.pdf. Acesso em: 10 abr. 2022.

SADOYAMA, A. et al. Evasão escolar no ensino superior: um estudo de revisão sistemática. **Revista de Psicologia**, Educação e Cultura, v. 24, n. 3, p. 92-103, 2020. Disponível em: <https://comum.rcaap.pt/handle/10400.26/34678>. Acesso em: 28 out. 2022.

SAGGI, M. K.; JAIN, S. A survey towards an integration of big data analytics to big insights for value-creation. **Information Processing & Management**, v. 54, n. 5, p. 758-790, 2018. Disponível em: <https://doi.org/10.1016/j.ipm.2018.01.010>. Acesso em: 20 ago. 2022.

SALDANHA, L. B.; CASTRO, S. F. As políticas de assistência estudantil em uma universidade federal brasileira: análise das políticas previstas no PNAES. **Revista Ciranda**, v. 5, n. 3, p. 336-352, 2021. Disponível em: <https://www.periodicos.unimontes.br/index.php/ciranda/article/view/4558>. Acesso em: 25 jun. 2022.

SALES, T. P.; GUIZZARDI, G. Ontological anti-patterns: Empirically uncovered error-prone structures in ontology-driven conceptual models. **Data & Knowledge Engineering**, v. 99, p. 72-104, 2015. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0169023X15000373>. Acesso em: 17 out. 2022.

SALES JUNIOR, J. S. et al. Análise estatística da evasão na Universidade Federal do Espírito Santo e uma avaliação de seus determinantes. **XLVII Simpósio Brasileiro de Pesquisa Operacional**, 2015. Disponível em: <http://cdsid.org.br/sbpo2015/wp-content/uploads/2015/08/142900.pdf>. Acesso em: 03 mai. 2022.

SALVINI, R. R. et al. Avaliação do impacto do Atendimento Educacional Especializado (AEE) sobre a defasagem escolar dos alunos da educação especial. **Estudos Econômicos (São Paulo)**, v. 49, p. 539-568, 2019. Disponível em: <https://www.scielo.br/j/ee/a/cPK5nWbDbfvn33T6tnqYYnh/?format=html&lang=pt>. Acesso em: 10 abr. 2024.

SANNA, F. A. **A importância da integração de dados para a qualidade da decisão do gestor público**. Trabalho de Conclusão de Curso. ENAP. Especialista em Gestão Pública. 2014. Disponível em: <https://repositorio.enap.gov.br/handle/1/1940>. Acesso em: 30 jan. 2023.

SANTOS, C. H.; MARTINS, S. L.; PLASTINO, A. É Possível Prever Evasão com Base Apenas no Desempenho Acadêmico?. **Simpósio Brasileiro de Informática na Educação**, v. 32, 2021. Disponível em: <https://sol.sbc.org.br/index.php/sbie/article/view/18107>. Acesso em: 12 mar. 2022.

SANTOS, C. O. **Ações institucionais para redução da evasão na UTFPR e no campus Medianeira**: entre o pactuado no REUNI e o concretizado. Tese de Doutorado. Universidade Tecnológica Federal do Paraná. Doutorado em Ensino de Ciência e Tecnologia. 2022. Disponível em: <http://repositorio.roca.utfpr.edu.br/jspui/handle/1/28966>. Acesso em: 8 jul. 2022.

SANTOS, C. P. C.; MARAFON, N. M. A política de assistência estudantil na universidade pública brasileira: desafios para o serviço social. **Textos & Contextos**, v. 15, n. 2, p. 408-422, 2016. Disponível em: https://www.researchgate.net/profile/Claudia-Santos-24/publication/313503065_A_Politica_d_e_Assistencia_Estudantil_na_Universidade_Publica_Brasileira_os_desafios_para_o_Servico_Social/links/589c82d692851c599c961271/A-Politica-de-Assistencia-Estudantil-na-Universidade-Publica-Brasileira-os-desafios-para-o-Servico-Social.pdf. Acesso em: 10 jun. 2022.

SANTOS, D. G.; SIQUEIRA, P. H. L.; DINIZ, D. M. Assistência Estudantil e PNAES: um estudo bibliométrico sobre a produção científica recente. **Práticas em Gestão Pública Universitária**, v. 5, n. 1, p. 28-53, 2021. Disponível em: <https://revistas.ufrj.br/index.php/pgpu/article/view/31056/23925>. Acesso em: 15 mar. 2022.

SANTOS, E. H. M. **As configurações da assistência estudantil no Brasil**: o cenário do PNAES. Dissertação de Mestrado. Universidade Federal de São Paulo. Mestrado em Serviço Social e Políticas Sociais. 2020. Disponível em: <http://200.144.93.16/bitstream/handle/11600/60792/Disserta%C3%A7%C3%A3o%20SANTO S.%20E.%20H.%20M..pdf?sequence=1&isAllowed=y>. Acesso em: 10 jul. 2022.

SANTOS, G.; FREITAS, L. O. Ensino superior público brasileiro: acesso e permanência no contexto de expansão. **Argumentum**, v. 6, n. 2, p. 182-200, jul./dez. 2014. Disponível em: <https://www.redalyc.org/pdf/4755/475547143014.pdf>. Acesso em: 18 mai. 2022.

SANTOS, J.; FIGUEIREDO, A. S.; VIEIRA, M. Rendimento acadêmico em ensino clínico e frequência às aulas: um estudo com estudantes de enfermagem. **Revista da Unidade de Investigação do Instituto Politécnico de Santarém**, v. 6, n. 2, p. 3-12, 2018. Disponível em: <https://revistas.rcaap.pt/uiips/article/view/16124>. Acesso em: 01 ago. 2022.

SANTOS, L. et al. Levantamento de fatores geradores de evasão no curso de Ciências Exatas - Licenciatura da UNIPAMPA. **Anais do Salão Internacional de Ensino**, Pesquisa e Extensão, v. 11, n. 2, 2019. Disponível em: <https://periodicos.unipampa.edu.br/index.php/SIEPE/article/view/101605>. Acesso em: 12 nov. 2022.

SANTOS, N. A. **Determinantes do desempenho acadêmico dos cursos de ciências contábeis**. Tese de Doutorado. Universidade de São Paulo. Doutorado em Ciências Contábeis. 2012. Disponível em: <https://www.teses.usp.br/teses/disponiveis/12/12136/tde-11062012-164530/pt-br.php>. Acesso em: 01 ago. 2022.

SAPUTRA, D. M.; SAPUTRA, D.; OSWARI, L. D. Effect of distance metrics in determining k-value in k-means clustering using elbow and silhouette method. **Sriwijaya international conference on information technology and its applications (SICONIAN 2019)**, p. 341-346, 2020. Disponível em: <https://www.atlantis-press.com/proceedings/siconian-19/125939938>. Acesso em: 19 abr. 2025.

SARACEVIC, T. **Interdisciplinary nature of information science**. 1995. Disponível em: <http://labds.eci.ufmg.br:8080/handle/123456789/75>. Acesso em: 18 jun. 2022.

SARAIVA, D. V. et al. Análise de Agrupamentos para Caracterização de Indicadores de Evasão. **WEI - Workshop sobre Educação em Computação**, n. 29, p. 238-247, 2021. Disponível em: <https://sol.sbc.org.br/index.php/wei/article/view/15915>. Acesso em: 04 ago. 2022.

SARAVIA, E.; FERRAREZI, E. **Políticas públicas: Coletâneas**. Brasília: ENAP, 2007.

SAVIC, D. Evolution of information resource management. **Journal of librarianship and information science**, v. 24, n. 3, p. 127-138, 1992. Disponível em: <https://doi.org/10.1177/096100069202400302>. Acesso em: 23 out. 2022.

SCIFO, E. **Graph Data Science with Neo4j**. Packt Publishing, 2023.

SECCHI, L. **Políticas públicas: conceitos, esquemas de análise, casos práticos**. São Paulo: Cengage Learning, 2014.

SENKEVICS, A. S.; MELLO, U. M. O perfil discente das universidades federais mudou pós-Lei de Cotas?. **Cadernos de Pesquisa**, v. 49, p. 184-208, 2019. Disponível em: <https://www.scielo.br/j/cp/a/KSvkm3DG3pPZYvpXxQc6PFh/?lang=pt>. Acesso em: 4 jul. 2022.

SHIRASU, M. R.; ARRAES, R. A. Determinantes da evasão e repetência escolar. **Encontro Nacional de Economia**, v. 43, 2016. Disponível em: <https://www.anpec.org.br/encontro/2015/submissao/files/I/i12-85f3c3774c3d65741cb278e01e61db39.pdf>. Acesso em: 05 mar. 2022.

SILVA, A. V.; FREITAS, T. A.; TEIXEIRA, G. Proposição de instrumento de avaliação de impacto do PNAES a partir de aplicação nos campi fora da sede da FURG. **Revista do Serviço Público**, v. 75, n. 2, p. 341-365, 2024. Disponível em: <https://repositorio.enap.gov.br/bitstream/1/7983/1/9079-Texto%20do%20Artigo-33686-1-10-20240627.pdf>. Acesso em: 17 abr. 2025.

SILVA, D. A. **Aplicação de técnicas de pré-processamento e agrupamento na base de dados de benefícios previdenciários do ministério público do trabalho.** Trabalho de Conclusão de Curso. Universidade Federal de Uberlândia. Graduação em Sistemas de Informação. 2018. Disponível em: <https://repositorio.ufu.br/handle/123456789/22118>. Acesso em: 25 ago. 2022.

SILVA, D. M. et al. Estilos de aprendizagem e desempenho acadêmico na Educação a Distância: uma investigação em cursos de especialização. **Revista Brasileira de Gestão de Negócios**, v. 17, p. 1300-1316, 2015. Disponível em: <https://www.scielo.br/j/rbgn/a/8pdVGdybJbSQmd5CShb6BmM/abstract/?lang=pt>. Acesso em: 17 jul. 2022.

SILVA, E. W. P.; MOREIRA, M. I. B. Inserção na Universidade Pública e Permanência Estudantil: Estudante na Corda Bamba. **Revista de Ensino, Educação e Ciências Humanas**, v. 23, n. 2, 2022. Disponível em: <https://revistaensinoeducacao.pgskroton.com.br/article/view/9284>. Acesso em: 9 jul. 2022.

SILVA, F. C. O desempenho acadêmico e o fenômeno da evasão em cursos de graduação da área da saúde. Dissertação de Mestrado. Universidade de Brasília. Mestrado em Ciências e Tecnologias em Saúde. 2016. Disponível em: <https://repositorio.unb.br/handle/10482/19947>. Acesso em: 22 abr. 2022.

SILVA, G. H. G. Ações afirmativas no ensino superior brasileiro: caminhos para a permanência e o progresso acadêmico de estudantes da área das ciências exatas. **Educação em Revista**, v. 35, 2019. Disponível em: <https://www.scielo.br/j/edur/a/JJds6YhTYpWn68F3XY3ccVS/abstract/?lang=pt>. Acesso em: 17 set. 2022.

SILVA, H. T.; RAMOS, L. F. M. Evasão de Curso: As Responsabilidades dos Agentes do Ensino Superior. **Revista UNIÍTALO em Pesquisa**, v. 8, n. 3, 2018. Disponível em: <http://pesquisa.italo.com.br/index.php?journal=uniitalo&page=article&op=view&path%5B%5D=227>. Acesso em: 10 mai. 2022.

SILVA, L. B.; COSTA, N. C. D. Acesso e permanência em desproporção: as insuficiências do programa nacional de assistência estudantil. **Agenda Política**, v. 6, n. 2, p. 166-192. 2018. Disponível em: <https://www.agendapolitica.ufscar.br/index.php/agendapolitica/article/view/208>. Acesso em: 12 set. 2022.

SILVA, M. A. **O Pré-Processamento em Mineração de Dados como método de suporte à modelagem algorítmica.** Dissertação de Mestrado. Universidade Federal do Tocantins. Mestrado em Modelagem Computacional de Sistemas. 2014. Disponível em: <https://docs.uft.edu.br/share/s/VbZ4G1KXQLqWbTuCPJqv7Q>. Acesso em: 23 ago. 2022.

SILVA, M. G. M.; NOGUEIRA, P. S. Permanência na Educação Superior: traçado dos estudos e faces da realidade. **Educação em Foco**, v. 18, n. 26, p. 43-68, 2015. Disponível em: <https://revista.uemg.br/index.php/educacaoemfoco/article/view/541>. Acesso em: 29 mai. 2022.

SILVA, R. M.; BOROWSKY, M. M. A implantação do Reuni na UFF: ampliação de direitos ou precarização do ensino superior?. **Regae-Revista de Gestão e Avaliação Educacional**, v. 7, n. 16, p. 91-110, 2018. Disponível em: <https://www.redalyc.org/journal/4718/471857006008/471857006008.pdf>. Acesso em: 15 jun. 2022.

SILVA, T. E.; TOMAÉL, M. I. A gestão da informação nas organizações. **Informação & Informação**, v. 12, n. 2, p. 148-149, 2007. Disponível em: <https://www.uel.br/revistas/uel/index.php/informacao/article/view/1806>. Acesso em: 30 set. 2022.

SILVA FILHO, R. L. L. et al. A evasão no ensino superior brasileiro. **Caderno de Pesquisa**, v. 37, n. 132, p. 641-659, 2007. Disponível em: <http://www.scielo.br/pdf/cp/v37n132/a0737132.pdf>. Acesso em: 2 mai. 2022.

SILVA FILHO, R. B.; ARAÚJO, R. M. L. Evasão e abandono escolar na educação básica no Brasil: fatores, causas e possíveis consequências. **Educação por escrito**, v. 8, n. 1, p. 35-48, 2017. Disponível em: <https://revistaseletronicas.pucrs.br/index.php/poescrito/article/view/24527>. Acesso em: 11 nov. 2022.

SILVEIRA, P. D. N. et al. Uma ontologia de referência para arquiteturas pedagógicas. **Anais do XXXII Simpósio Brasileiro de Informática na Educação**, p. 24-34, 2021. Disponível em: <https://sol.sbc.org.br/index.php/sbie/article/view/18040>. Acesso em: 04 jan. 2023.

SMITH, B.; BROCHHAUSEN, M. Putting biomedical ontologies to work. **Methods of information in medicine**, v. 49, n. 02, p. 135-140, 2010. Disponível em: <https://www.thieme-connect.com/products/ejournals/abstract/10.3414/ME9302>. Acesso em: 29 set. 2022.

SOARES, T. C. S. A Reestruturação e desestruturação da política de ensino superior no Brasil: os avanços e retrocessos na contradição expansão com contenção. **Revista Serviço Social em Perspectiva**, v. 6, n. 01, p. 166-187, 2022. Disponível em: <https://www.periodicos.unimontes.br/index.php/sesoperspectiva/article/view/4680>. Acesso em: 13 jul. 2022.

SOUSA, A. P. L. D. **Inteligência emocional e desempenho acadêmico em estudantes do ensino superior**. Dissertação de Mestrado. ISMT. Mestrado em Gestão de Recursos Humanos e Comportamento Organizacional. 2010. Disponível em: <https://dspace.ismt.pt/bitstream/123456789/291/1/Tese%20de%20Mestrado%20-%20Intelig%c3%aancia%20Emocional%20e%20Desempenho%20Acad%c3%a9mico%20em%20Estudantes%20antes%20do%20Ensino%20Superior.pdf>. Acesso em: 18 jul. 2022.

SOUSA, A. S.; OLIVEIRA, G. S.; ALVES, L. H. A pesquisa bibliográfica: princípios e fundamentos. **Cadernos da FUCAMP**, v. 20, n. 43, 2021. Disponível em: <https://revistas.fucamp.edu.br/index.php/cadernos/article/view/2336>. Acesso em: 23 nov. 2022.

SOUSA, M. C. C. **Uma análise do algoritmo K-means como introdução ao aprendizado de máquinas**. Trabalho de Conclusão de Curso. Universidade Federal do Tocantins. Licenciatura em Matemática. 2020. Disponível em: <http://repositorio.uft.edu.br/handle/11612/1764>. Acesso em: 26 ago. 2022.

SOUSA, Z. A. S.; FERREIRA, M. A.; MIRANDA, G. J. Teoria da atribuição de causalidade: percepções docentes e discentes sobre os determinantes do desempenho acadêmico. **Advances in Scientific & Applied Accounting**, v. 12, n. 2, 2019. Disponível em: https://www.researchgate.net/profile/Gilberto-Miranda/publication/337335674_TEORIA_DA_ATRIBUICAO_DE_CAUSALIDADE_PERCEPCOES_DOCENTES_E_DISCENTES SOBRE_OS_DETERMINANTES_DO_DESEMPENHO_ACADEMICO/links/5dd32594a6fdcc7e138d29ed/TEORIA-DA-ATRIBUICAO-DE-CAUSALIDADE-PERCEPCOES-DOCENTES-E-DISCE

[NTES-SOBRE-OS-DETERMINANTES-DO-DESEMPENHO-ACADEMICO.pdf](#). Acesso em: 19 jul 2022.

SOUZA, E. F.; FALBO, R. A.; VIJAYKUMAR, N. L. ROoST: Reference ontology on software testing. **Applied Ontology**, v. 12, n. 1, p. 59-90, 2017. Disponível em: <https://content.iospress.com/articles/applied-ontology/ao177>. Acesso em: 04 jan. 2023.

SOUZA, J. M.; TOASSI, R. F. C. Evasão na Educação Superior: estudo de caso do curso noturno de Odontologia da Universidade Federal do Rio Grande do Sul. **Colóquio Internacional de Gestão Universitária - CIGU**, v. 14, 2014. Disponível em: <https://lume.ufrgs.br/handle/10183/108503>. Acesso em: 04 set. 2022.

SOUZA, M. P. N. Abordagem inter e transdisciplinar em ciência da informação. In: TOUTAIN, L. M. B. B. **Para entender a ciência da informação**. Salvador: EDUFBA, 2012. Disponível em: <https://repositorio.ufba.br/bitstream/ufba/145/1/Para%20entender%20a%20ciencia%20da%20informacao.pdf>. Acesso em: 15 mai. 2022.

SOUZA, R. O. C. **Comportamento informacional dos gestores de assuntos estudantis das universidades públicas brasileiras**. Dissertação de Mestrado. Universidade Federal do Paraná. Mestrado em Gestão da Informação. Disponível em: <https://www.acervodigital.ufpr.br/handle/1884/67594>. Acesso em: 20 fev. 2022.

SOUZA, W. V. B. **Enfrentamento à COVID-19 nas universidades federais brasileiras: uma pesquisa-ação no Ministério da Educação**. Tese de Doutorado. Universidade de Brasília. Doutorado em Administração. 2021. Disponível em: <https://repositorio.unb.br/handle/10482/41986>. Acesso em: 29 set. 2022.

SUTRAMIANI, N. P. et al. The Performance Comparison of DBSCAN and K-Means Clustering for MSMEs Grouping based on Asset Value and Turnover. **Journal of Information Systems Engineering & Business Intelligence**, v. 10, n. 1, 2024. Disponível em: <https://e-journal.unair.ac.id/JISEBI/article/download/47770/28047>. Acesso em: 15 abr. 2025.

SVEIBY, K. E. **A Nova Riqueza das Organizações**. Rio de Janeiro: Campus, 1998.

TABOSA, H. R.; PINTO, V. B. Caracterização do comportamento de busca e uso de informação na área da saúde: o modelo de Ellis aplicado ao estudo do comportamento informacional de pacientes. **Informação & Sociedade: Estudos**, v. 26, n. 2, p. 225-237, 2016. Disponível em: <https://repositorio.ufc.br/handle/riufc/24462>. Acesso em: 28 out. 2022.

TAN, P. N. et al. **Introduction to Data Mining**. São Paulo: Pearson, 2019.

TARAPANOFF, K. **Aprendizagem Organizacional: fundamentos e abordagens multidisciplinares**. Curitiba: IBEPEX, 2011.

TARAPANOFF, K. **Inteligência organizacional e competitiva**. Brasília: Editora Universidade de Brasília, 2001.

TELES, J. et al. Desempenho acadêmico dos estudantes: uma análise dos fatores preditivos. **FACEF Pesquisa-Desenvolvimento e Gestão**, v. 22, n. 1, 2019. Disponível em: <http://periodicos.unifacef.com.br/index.php/facefpesquisa/article/view/1577>. Acesso em: 11 jul. 2022.

TREVIZAN, E. A. **Programa de apoio a planos de reestruturação e expansão das universidades federais (Reuni): contextos, condicionantes e resultados de sua**

implementação. Tese de Doutorado. Universidade Estadual Paulista. Doutorado em Educação. 2019. Disponível em: <https://repositorio.unesp.br/handle/11449/191147>. Acesso em: 25 mai. 2022.

TODD, R. J. Adolescents of the information age: patterns of information seeking and use, and implications for information professionals. **School libraries worldwide**, v. 9, n. 2, p. 27-46, 2003. Disponível em: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=a66e830699a61138d36f66d5688944244786be45>. Acesso em: 02 set. 2022.

TOSIN, T. M. **Um modelo para integração de informações de bases de dados abertos, com uso de ontologias**. Dissertação de Mestrado. Unisinos. Mestrado em Computação Aplicada. 2016. Disponível em: <http://www.repositorio.jesuita.org.br/handle/UNISINOS/5244>. Acesso em: 15 jan. 2023.

TREVISOL, J. V.; NIEROTKA, R. L. Lei das cotas e as políticas de democratização do acesso ao ensino superior público brasileiro. **Quaestio - Revista de Estudos em Educação**, v. 17, n. 2, 2015. Disponível em: <http://periodicos.uniso.br/ojs/index.php/quaestio/article/view/2406>. Acesso em: 11 jul. 2022.

USCHOLD, M.; GRUNINGER, M. Ontologies: Principles, methods and applications. **The knowledge engineering review**, v. 11, n. 2, p. 93-136, 1996. Disponível em: <https://www.cambridge.org/core/journals/knowledge-engineering-review/article/abs/ontology-s-principles-methods-and-applications/2443E0A8E5D81A144D8C611EF20043E6>. Acesso em: 25 set. 2022.

USCHOLD, M.; JASPER, R. A framework for understanding and classifying ontology applications. **Proceedings of the IJCAI-99 workshop on ontologies and Problem-Solving Methods (KRR5)**, 1999. Disponível em: <https://people.cs.ksu.edu/~abreed/CIS890/References/uschold99.pdf>. Acesso em: 21 mar. 2023.

VALENTIM, M. L. P. Inteligência competitiva em organizações: dado, informação e conhecimento. **DataGramZero**, Rio de Janeiro, v. 3, n. 4, p. 1-13, 2002. Disponível em: https://www.brapci.inf.br/_repositorio/2010/01/pdf_f589d25523_0007468.pdf. Acesso em: 18 out. 2022.

VASCONCELOS, N. B. Programa Nacional de Assistência Estudantil: uma análise da evolução da assistência estudantil ao longo da história da educação superior no Brasil. **Ensino Em-Revista**, v. 17, n. 2, p. 599-616, jul./dez. 2010. Disponível em: <http://www.seer.ufu.br/index.php/emrevista/article/view/11361/6598>. Acesso em: 2 jul. 2022.

VAZQUEZ, D.; DELAPLACE, D. Políticas públicas na perspectiva de direitos humanos: um campo em construção. **Revista internacional de direitos humanos**, v. 8, n. 14, p. 34-65, 2011. Disponível em: <https://core.ac.uk/download/pdf/16031669.pdf>. Acesso em: 11 abr. 2024.

VELOSO, C. M. L.; COUTO, A. C. S. R.; VALENTIM, M. C. O nivelamento escolar como instrumento de redução da evasão no curso de Engenharia Civil-FACEMG. **The Journal of Engineering and Exact Sciences**, v. 4, n. 4, p. 0399-0404, 2018. Disponível em: <https://periodicos.ufv.br/jcec/article/view/2566>. Acesso em: 10 out. 2022.

VIANA, A. L. Abordagens metodológicas em políticas públicas. **Revista De Administração Pública**, v. 30, n. 2, p. 5-43, 1996. Disponível em: <https://periodicos.fgv.br/rap/article/view/8095>. Acesso em: 29 abr. 2024.

VITELLI, R. F.; FRITSCH, R. Evasão escolar na educação superior: de que indicador estamos falando?. **Estudos em Avaliação Educacional**, v. 27, n. 66, p. 908-937, 2016. Disponível em: <https://publicacoes.fcc.org.br/eae/article/view/4009>. Acesso em: 30 set. 2022.

WANG, H. et al. Feature selection strategies: a comparative analysis of SHAP-value and importance-based methods. **Journal of Big Data**, v. 11, n. 1, p. 44, 2024. Disponível em: <https://link.springer.com/article/10.1186/s40537-024-00905-w>. Acesso em: 15 abr. 2025.

WANG, Z.; TAYLOR, M. E. Improving Reinforcement Learning with Confidence-Based Demonstrations. **IJCAI**, p. 3027-3033, 2017. Disponível em: <https://www.ijcai.org/proceedings/2017/0422.pdf>. Acesso em: 30 ago. 2022.

WILSON, T. D. Information Behaviour: an interdisciplinary perspective. University of Sheffield Department of Information Studies, **British Library Research and Innovation Report**, n. 10. 1996. Disponível em: <https://webcitation.org/5SuU1VliM>. Acesso em: 01 fev. 2023.

WILSON, T. D. Information Management. **International Encyclopedia of Information and Library Science**, Routledge, London, p. 263-278, 2002. Disponível em: http://informationr.net/tdw/publ/papers/encyclopedia_entry.html. Acesso em: 18 out. 2022.

WILSON, T. D. Models in information behaviour research. **Journal of documentation**, v. 55, n. 3, p. 249-270, 1999. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/EUM0000000007145/full/html>. Acesso em: 15 abr. 2022.

ZAGO, N. Do acesso à permanência no ensino superior: percursos de estudantes universitários de camadas populares. **Revista Brasileira de Educação**, v. 11, p. 226-237, 2006. Disponível em: <https://www.scielo.br/rbedu/a/wVchYRqNFkssn9WqQbj9sSG/?lang=pt>. Acesso em: 1 jul. 2022.

APÊNDICE A – PESQUISA SOBRE O COMPORTAMENTO INFORMACIONAL

Pesquisa sobre o Comportamento Informacional dos servidores da Assistência Estudantil do IFMG

Prezado(a) Servidor(a),

Esta pesquisa está sendo aplicada com o objetivo de identificar características do Comportamento Informacional dos servidores que atuam na Assistência Estudantil dos campi do Instituto Federal de Minas Gerais (IFMG). Trata-se de uma parte do estudo previsto para a elaboração de uma tese de Doutorado do Programa de Pós-Graduação em Gestão & Organização do Conhecimento da Universidade Federal de Minas Gerais (UFMG).

De acordo com nossas estimativas, o preenchimento do questionário não leva mais do que cinco minutos. Na próxima página será apresentado o Termo de Consentimento Livre e Esclarecido para sua leitura e concordância.

É importante ressaltar que as informações fornecidas são sigilosas em relação à identificação do respondente e utilizadas para fins exclusivamente científicos.

Desde já, nossos agradecimentos pela participação!

Atenciosamente,

Eduardo Cardoso Melo (Doutorando)

Marcello Peixoto Bax (Orientador e Professor da UFMG)

Termo de Consentimento Livre e Esclarecido

Você está sendo convidado (a) como voluntário(a) a participar da pesquisa sobre o Comportamento Informacional dos servidores da Assistência Estudantil do IFMG no contexto de um estudo do Programa de Pós-Graduação em Gestão & Organização do Conhecimento da Universidade Federal de Minas Gerais (UFMG).

Nesta etapa o procedimento de coleta de dados será por meio de um questionário.

FORMA DE ACOMPANHAMENTO E ASSISTÊNCIA E BENEFÍCIOS: Você será acompanhado pelo responsável pela pesquisa em qualquer momento que se fizer necessário, podendo enviar um e-mail ao mesmo. Os benefícios da pesquisa serão indiretos, podendo trazer informações sobre as práticas de uso da informação no ambiente de trabalho da Assistência Estudantil do IFMG.

ARMAZENAMENTO DE DADOS: Ao final da pesquisa, todo o material será mantido permanentemente em um banco de dados de pesquisa, de acesso restrito, sob a responsabilidade do pesquisador.

GARANTIA DE ESCLARECIMENTO, LIBERDADE DE RECUSA E GARANTIA DE SIGILO: Você será esclarecido sobre a pesquisa em qualquer aspecto que desejar. Você é livre para recusar-se a participar, retirar seu consentimento ou interromper a participação a qualquer momento. A sua participação é voluntária e a recusa ou desistência em participar não acarretará qualquer penalidade ou perda de benefícios. O pesquisador irá tratar a sua identidade com padrões profissionais de sigilo. Você não será identificado em nenhuma publicação que possa resultar deste estudo. Uma via deste consentimento será arquivada com os pesquisadores e outra poderá ser fornecida a você.

CUSTOS DA PARTICIPAÇÃO, RESSARCIMENTO E INDENIZAÇÃO POR EVENTUAIS DANOS: A participação no estudo não acarretará custos para você e não será disponível nenhuma compensação financeira adicional. Se houver algum dano, decorrente da pesquisa, você terá direito a buscar indenização, por meio das vias judiciais.

Declaro que fui informado(a) dos objetivos da pesquisa de maneira clara e detalhada e esclareci minhas dúvidas. Sei que em qualquer momento poderei receber ou solicitar novas informações por e-mail.

O pesquisador Eduardo Cardoso Melo certifica-me de que todos os dados desta pesquisa são confidenciais. Também sei que caso existam gastos adicionais, estes serão absorvidos pelo orçamento da pesquisa. Em caso de dúvidas, exclusivamente relativas ao projeto, poderei chamar o pesquisador citado nos contatos:

Eduardo Cardoso Melo

Fone: (XX) XXXXX-XXXX

E-mail: eduardo.melo@ifmg.edu.br

Ao clicar em Concordo, você informa que está de acordo com este Termo de Consentimento Livre e Esclarecido e que deseja prosseguir para o questionário, nesta data.

Você concorda em participar desta pesquisa?

Sim

Não

Parte 1 - Caracterização do respondente

1.1) Identidade de gênero

- Feminino
- Masculino
- Outra

1.2) Campus onde atua

1.3) Idade atual

1.4) Há quanto tempo trabalha no Setor de Assistência Estudantil?

- Até 5 anos
- Entre 6 e 10 anos
- Entre 11 e 15 anos
- Entre 16 e 20 anos
- Acima de 20 anos

Parte 2 - Comportamento Informacional

2.1) Como você reconhece ou identifica a necessidade de determinada informação:

LEGENDA: (1) Nunca (2) Raramente (3) Ocasionalmente (4) Frequente (5) Muito frequente (6) Não se aplica						
Necessidade de informação	1	2	3	4	5	6
Sanar dúvidas						
Tomada de decisão						
Identificação de uma lacuna						
Construção do conhecimento						

2.2) Com qual intensidade você utilizou informações da Instituição na qual exerce suas funções sobre... ?

LEGENDA: (1) Nunca (2) Raramente (3) Ocasionalmente (4) Frequente (5) Muito frequente (6) Não se aplica						
Tipo de informação	1	2	3	4	5	6
Perfil socioeconômico do estudante						
Estudantes com necessidades educacionais específicas						
Acompanhamento pedagógico e/ou social dos estudantes						
Aproveitamento acadêmico dos estudantes						
Evasão estudantil						
Refeições servidas pelo Restaurante Universitário						
Orçamento e Finanças						
Processos Licitatórios						
Contratação de pessoal						
Capacitação dos servidores						

2.3) Os itens analisados na questão anterior (2.2) foram utilizados por você para qual propósito? Marque sua resposta na escala de acordo com a frequência de utilização do propósito.

LEGENDA: (1) Nunca (2) Raramente (3) Ocasionalmente (4) Frequente (5) Muito frequente (6) Não se aplica						
Tipo de propósito	1	2	3	4	5	6
Elaboração de ações/metast para assistência estudantil						
Elaboração de políticas afirmativas e de acessibilidade						
Elaboração de relatórios para prestação de contas						
Tomada de decisão						
Aquisição de bens e materiais						
Composição da equipe de trabalho						
Proposição de debates sobre temas específicos da área						
Construção de parcerias com outras instituições						
Proposição de cursos e oficinas para estudantes e servidores						

2.4) Durante o processo de busca da informação, você identificou algum tipo de barreira? Com qual frequência?

LEGENDA: (1) Nunca (2) Raramente (3) Ocasionalmente (4) Frequente (5) Muito frequente (6) Não se aplica						
Barreiras encontradas	1	2	3	4	5	6
Tempo						
Idioma						
Geográfica						
Pessoal qualificado ou disponível						
Falta de um sistema informacional específico						
Manuseio de ferramentas tecnológicas disponíveis						

2.5) Durante o processo de busca por uma informação, aponte se foram identificados alguns desses sentimentos e com qual frequência?

LEGENDA: (1) Nunca (2) Raramente (3) Ocasionalmente (4) Frequente (5) Muito frequente (6) Não se aplica						
Sentimentos	1	2	3	4	5	6
Angústia						
Curiosidade						
Medo						
Estresse						
Aflição						
Incerteza						
Apreensão						
Ansiedade						

2.6) Com que frequência você utilizou as seguintes fontes de informações? (As informações foram provenientes de...)

LEGENDA: (1) Nunca (2) Raramente (3) Ocasionalmente (4) Frequente (5) Muito frequente (6) Não se aplica						
Fontes de informação	1	2	3	4	5	6
Fontes internas:						
SEI / SUAP / SSAE						
Ofícios e memorandos						
Fóruns/Grupos de aplicativos						
Colegas de trabalho						
Fontes externas:						
Prefeitura Municipal / Governo Estadual						
Governo Federal (ex: IBGE)						
Ministério Público / Tribunal de Contas						
MEC / Plataformas de indicadores (ex: Nilo Peçanha)						
IBGE						
Sindicatos						

2.7) Com que intensidade você utilizou os seguintes canais de informação (As informações chegaram até você ou você as procurou utilizando quais veículos)?

LEGENDA: (1) Nunca (2) Raramente (3) Ocasionalmente (4) Frequente (5) Muito frequente (6) Não se aplica						
Canais de informação	1	2	3	4	5	6
Contato pessoal						
Telefone						
Sistemas de Informação						
Sites institucionais						
E-mail						
Redes sociais						
Portal da Transparência						
Sites de pesquisa (ex: Google)						
Jornais, revistas ou livros						

2.8) Como você costuma compartilhar as informações? Em qual intensidade?

LEGENDA: (1) Nunca (2) Raramente (3) Ocasionalmente (4) Frequente (5) Muito frequente (6) Não se aplica						
Compartilhamento de informação	1	2	3	4	5	6
Contato pessoal						
Telefone						
Sistemas de Informação						
Sites institucionais						
Redes sociais						
E-mail						
Aplicativos de mensagem (Whatsapp, Telegram, etc)						

2.9) Quais desses itens você considera ao compartilhar informações? Com qual frequência?

LEGENDA: (1) Nunca (2) Raramente (3) Ocasionalmente (4) Frequente (5) Muito frequente (6) Não se aplica						
Compartilhamento de informação	1	2	3	4	5	6
Confiança						
Risco						
Proximidade						
Medo						
Insegurança						

Parte 3 - Informação valiosa

3.1) Os itens seguintes correspondem às características gerais de uma "boa informação" ou "informação valiosa". Avalie cada um deles com a importância que você lhes atribui quando relacionados às informações necessárias para a execução dos trabalhos no âmbito da assistência estudantil.

Características da informação	Mais importante	Menos importante
Precisa (não contém erro)		
Completa (contém todos os fatos relevantes)		
Econômica (não demandou custo excessivo na sua disponibilização)		
Flexível (poderá ser usada para uma variedade de propósitos)		
Confiável (os métodos de coleta e a origem são conhecidos)		
Relevante (essencial para o processo de tomada de decisões)		
Simple (de fácil entendimento, sem muitos detalhes)		
Pontual (disponibilizada dentro do prazo estipulado)		
Verificável (possível de se conferir e assegurar que está correta)		
Acessível (de fácil obtenção por parte dos usuários autorizados)		
Segura (o acesso só é permitido aos usuários autorizados)		