



XXII Encontro Nacional de Pesquisa em Ciência da Informação – XXII ENANCIB

ISSN 2177-3688

GT-11 – Informação & Saúde

USO DE DICIONÁRIO SEMÂNTICOS NA ESTRATIFICAÇÃO DE RISCOS EM SAÚDE MENTAL

USE OF SEMANTIC DATA DICTIONARY IN THE STRATIFICATION OF RISK MENTAL DISORDERS

Evaldo de Oliveira da Silva. UFMG.

Marcello Peixoto Bax. UFMG.

Modalidade: Trabalho Completo

Resumo: Equipes assistenciais utilizam a estratificação de riscos na área da saúde mental para determinar o nível de risco e a gravidade dos sinais e sintomas apresentados pelos pacientes. Os dados coletados são estruturados em grupos de questões. Esta estrutura não favorece a análise dos resultados com base nas questões respondidas. As equipes assistenciais necessitam que os dados estejam organizados e preparados para responder a diferentes questões, ou ainda, sobre a gravidade dos sinais e sintomas estratificados. Ontologias têm sido utilizadas para organizar e descrever o conhecimento sobre dados. Facilitam a extração de informações, desambiguam termos e preservam a semântica das variáveis da saúde do paciente usando anotação semântica de dados. Assim, “como melhorar a análise dos dados sobre a estratificação de riscos em saúde mental a fim de orientar a definição dos níveis de cuidados em saúde mental?” Para responder a esta questão propõe-se a modelagem ontológica da estratificação de riscos e anotação semântica dos dados utilizando a técnica *Semantic Data Dictionary* (SDD). O SDD utiliza dicionários semânticos de dados para gerar grafos de conhecimento a partir de ontologias e de templates de metadados. Os grafos podem ser explorados para filtrar dados necessários para verificar hipóteses em diferentes dados na área da saúde mental. Como resultado da pesquisa, grafos são gerados para facilitar a visualização das informações e a descoberta de relações entre, p.ex., doenças, sinais e sintomas, servindo como ferramenta para aprimorar o processo de estratificação, além de auxiliar na definição dos níveis de cuidados com o paciente.

Palavras-Chave: Estratificação. Saúde Mental. Ontologia. Anotação Semântica.

Abstract: Care teams use risk stratification in mental health to determine the level of risk and the severity of the signs and symptoms presented by patients. Risk stratification allows extracting data which are structured into groups of questions. This structure does not favor the analysis of the results based on the answered questions. Health care teams need data to be organized and prepared to answer different questions, or even the severity of signs and symptoms stratified. Ontologies have been used to organize and describe knowledge about data collections. They facilitate the extraction of information, disambiguate terms, and preserve the semantics of the patient's health variables using semantic data annotation. Thus, "how to improve the analysis of data on risk stratification in mental health in order to guide the definition of levels of mental health care?" To answer this question, we propose ontological modeling of risk stratification and semantic annotation of data using the SDD (*Semantic Data Dictionary*) technique. SDD uses semantic data dictionaries to generate knowledge graphs from ontologies and metadata templates. As a result of the research, graphs are generated to facilitate the visualization of information and the discovery of relationships between, e.g., diseases,

signs, and symptoms, serving as a tool to improve the stratification process, as well as helping to define levels of patient care.

Keywords: Stratification. Mental health. Ontology. Semantic Annotation.

1 INTRODUÇÃO

Na área da saúde, estratificar é reconhecer os diferentes graus de riscos de cada paciente com determinado agravo (BRASIL, 2013; ESPÍRITO SANTO, 2018). A estratificação de risco melhora o trabalho das equipes multiprofissionais, identificando os usuários do sistema de saúde por níveis de riscos. Contribui para o planejamento da assistência com plano de cuidados de acordo com a realidade de cada paciente (MELLO, 2015).

Equipes multiprofissionais na área da saúde mental utilizam a estratificação de riscos para determinar a gravidade dos sintomas apresentados. Grupos de questões no formato de questionário estruturam as coleções de dados (*datasets*) da estratificação submetida aos pacientes (PARANÁ, 2021; PAULA, 2019). No entanto, essa estrutura não favorece a análise de diferentes questões a serem respondidas sobre os pacientes. É necessário que os dados estejam preparados e organizados para auxiliar na definição dos diferentes parâmetros de assistência à saúde e elaboração do plano de cuidados, sem interpretações equivocadas dos diferentes termos que definem os dados extraídos da estratificação de riscos.

Ontologias organizam o conhecimento sobre coleções de dados. Na área da saúde as ontologias facilitam a extração de informações, desambigam termos e preservam a semântica das variáveis de saúde, tornando possível a anotação semântica de dados produzidos por diferentes estudos (BAX e SILVA, 2020; CEUSTERS e SMITH, 2010; HASTINGS, 2012; YAMADA et al, 2020). Desta forma, “como melhorar a preparação dos dados sobre a estratificação de riscos em saúde mental a fim de orientar análises sobre os planos de cuidados em saúde mental?” Propõe-se aplicar um processo sistemático de integração de dados com base na modelagem ontológica da estratificação de riscos e anotação semântica dos dados utilizando a técnica *Semantic Data Dictionary* (SDD¹) (RASHID et. al, 2017). A técnica SDD utiliza dicionários semânticos de dados para gerar grafos a partir de ontologias e de templates de metadados. Os grafos podem integrar dados oriundos de diferentes pesquisas e ser explorados posteriormente para filtrar dados necessários para verificar hipóteses em estudos na área da saúde mental. Podem ser úteis para encontrar relações entre doenças e

¹ Disponível em: <https://github.com/tetherless-world/SemanticDataDictionary>

sintomas, além de auxiliar na definição dos planos de cuidados do paciente. O uso do SDD é mais acessível ao especialista de domínio, já que as abordagens alternativas exigem conhecimentos técnicos em linguagens de marcação tais como JSON² ou RDF³ ou linguagens de programação. O SDD permite sistematizar o processo de conversão do *dataset* para o grafo pelo próprio especialista de domínio, uma vez que os dicionários são semelhantes aos tradicionais dicionários de dados, construídos por aqueles que publicam dados (RASHID et al., 2020).

Este trabalho apresenta um “caso de uso” do SDD no domínio de estratificação de riscos de saúde mental. O objetivo é preparar e organizar os dados deste domínio para potencializar análises futuras. Além disso, vale ressaltar, que esta abordagem se situa nas fases iniciais do ciclo de vida do *data science*, que preconiza a organização conceitual, na preparação de dados em ciência de dados (MICROSOFT, 2020).

No restante do artigo a Seção 2 descreve os conceitos da estratificação de riscos na saúde mental, plano de cuidados, ontologias na saúde mental e grafos de conhecimento. A Seção 3 apresenta a técnica SDD e como pode ser aplicada. A Seção 4 traz um estudo de caso do uso de SDD para estratificação de riscos em saúde mental. A Seção 5 correlaciona outros trabalhos encontrados na literatura. Ao final, na Seção 6, são apresentadas as conclusões.

2 ORGANIZAÇÃO DA INFORMAÇÃO EM SAÚDE MENTAL

2.1 Estratificação de Riscos em Saúde Mental

A estratificação de risco é um processo estatístico que apoia o tratamento do paciente ajustado ao risco atribuído. Miller, Reardon e Safi (2001) abordam a importância da coleta rigorosa de dados da estratificação de risco sobre dados clínicos, com base na elaboração de questões, que apoiam a comparação de dados de doenças ou combinação de tratamento-doença. Os autores destacam a necessidade das variáveis da estratificação serem identificadas cuidadosamente para evitar interpretações ambíguas.

A estratificação de risco é aplicada a partir da gravidade dos sinais e sintomas apresentados, sem a necessidade de firmar diagnóstico inicial. É um instrumento que oferece suporte às equipes de atenção primária, e estabelece estratégias de atendimento do que se

² JavaScript Object Notation. Disponível em: <https://www.json.org/json-en.html>

³ Resource Description Framework. Especificação disponível em: <https://www.w3.org/TR/rdf11-concepts/>

denomina “condições crônicas de saúde” (BRASIL, 2011; PARANÁ, 2021). A SESA⁴ aplica a estratificação de risco em saúde mental com base em grupos de transtornos mentais nomeados e classificados pela OMS⁵. A partir destes grupos, enumeram-se os principais sintomas psicopatológicos, fatores agravantes e atenuantes que podem ser identificados a partir da presença das principais síndromes psicopatológicas seguindo os critérios diagnósticos da CID 10 e do DSM IV-R (AMERICAN PSYCHIATRIC ASSOCIATION, 1994; PARANÁ, 2021; WHO, 1993). Perguntas submetidas aos pacientes estruturam os grupos de sinais e sintomas, de acordo com a frequência em que se apresentam nas respectivas síndromes psicopatológicas, e são pontuados de acordo com o grau de gravidade.

Paula (2019) analisa as classificações de risco como protocolos para dimensionar os cuidados aos usuários. Aplica técnicas para a coleta de dados usando o Questionário de Estratificação de Risco em Saúde Mental (QERSM). O tratamento dos dados é realizado por meio da categorização e análise temática após transcrição das entrevistas.

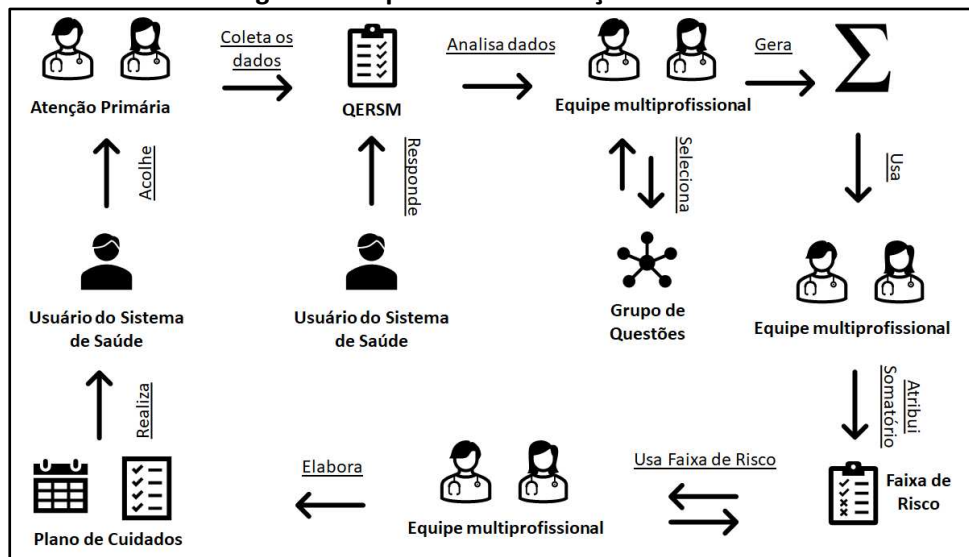
O projeto TelePSI propõe um guia que orienta a estratificação de riscos em saúde mental (BRASIL, 2020). Considera situações de risco em psiquiatria: risco de suicídio, risco de heteroagressão e risco de exposição moral. Essas situações apoiam a estratificação dos riscos em níveis: baixo, moderado ou risco (MOTTA et al., 2020). Riscos baixos a moderados orientam intervenções breves e encaminhamentos para guiar os profissionais no planejamento de cuidados. De acordo com o *National Confidential Inquiry into Suicide and Safety in Mental Health* (NCISH) a avaliação de risco combina fatores psicológicos e sociais para dimensionar os planos de cuidados dos pacientes. Iniciativas do NCISH envolve a utilização do preenchimento de um grupo de questões, disponibilizados pelo portal da instituição (NCISH, 2018).

Um *Mental Health Information Systems* (MHIS) visa orientar o atendimento à população acometida por transtornos mentais, apoiando as decisões sobre a qualidade do atendimento contemplando etapas para coleta de dados, agrupamento e classificação, análise, divulgação e uso de dados sobre diferentes estudos (OMS, 2008). A estratificação de risco na saúde mental obedece a um processo que é iniciado a partir do acolhimento do paciente no sistema de saúde, utilizando de instrumentos para coleta de dados, processamento e análise das informações. A Figura 1 resume este processo.

⁴ Secretaria de Saúde do Estado do Paraná

⁵ Organização Mundial da Saúde

Figura 1: Etapas da estratificação de riscos.



Fonte: Elaborado pelo autor.

2.2 Projeto Terapêutico Singular

De acordo com o Ministério da Saúde, a Atenção Básica deve cumprir funções para contribuir com o funcionamento das Redes de Atenção à Saúde, tais como: identificação de riscos, necessidades e demandas de saúde e intervenções clínicas e sanitárias efetivas. A implementação destas funções amplia os graus de autonomia dos indivíduos e permite coordenar o cuidado com o usuário do sistema de saúde, por meio de projetos terapêuticos singulares (PTS) (BRASIL, 2013). A criação de um PTS é uma ação que visa diminuir a recorrência de doenças mentais ou riscos do paciente com doença mental, seguindo também as orientações da OMS. O PTS deve ser elaborado a partir da estratificação de risco como um conjunto de atividades e de condutas terapêuticas articuladas para o paciente. A sistematização dos dados sobre a estratificação de riscos, e, como consequência, a geração de dados para o PTS torna-se importante para o trabalho de acompanhamento do paciente, a fim de reduzir a incidência de transtornos mentais da população (PARANÁ, 2021; SAXENA e SETOYA, 2013).

A representação do conhecimento sobre a estratificação de riscos pode enriquecer o entendimento sobre o domínio, sistematizando os diferentes dados utilizados. Ontologias enriquecem de semântica os dados extraídos da estratificação, a fim de representar o conhecimento sobre o nível em que ocorrerá a assistência em saúde mental. A próxima seção descreve como as ontologias na área da saúde servem de mecanismos de modelagem conceitual e anotação dos dados coletados.

2.3 Ontologias de Saúde Mental

Ceusters e Smith (2010) apresentam uma visão da doença mental baseada no realismo ontológico e que segue os princípios incorporados na *Basic Formal Ontology* (BFO) e na aplicação da BFO na *Ontology of General Medical Science* (OGMS). Os autores analisaram afirmações sobre o que é considerado como doença mental no contexto da DSM-V. Avaliam se a representação do conhecimento presente nas ontologias BFO e OGMS são adequadas para criar uma ontologia da saúde mental necessária para evitar interpretações erradas sobre um mesmo conceito. Brenas, Shin e Shaban-Nejad (2019) implementaram uma ontologia formal sobre experiências adversas da infância (ACEs, *Adverse Childhood Experiences*), para integração de dados de estudos científicos na área da saúde mental. Segundo os autores, as ACEs estão relacionadas aos riscos aumentados a partir dos resultados e condições de saúde dos pacientes. De acordo com Hastings (2012) as ontologias podem ser usadas para aplicações sofisticadas de raciocínio automatizado. Os autores descrevem a Ontologia do Funcionamento Mental (MF) e a Ontologia da Doença Mental (MD), duas ontologias baseadas em realismo para a descrição do funcionamento mental humano e das doenças. Benfares, Idrissi e Hamid (2018) propõe uma arquitetura de sistema de acompanhamento de pacientes, que proporciona o acompanhamento do estado psíquico do paciente. Os autores usam ontologias e sistemas de recomendação para proporcionar aos pacientes o acompanhamento em tempo real. Foram desenvolvidas rotinas para recuperar informações de rastreamento dos pacientes com câncer, depressão e transtornos de ansiedade. A pesquisa foi realizada na unidade de psicologia, no centro de oncologia e hematologia do Centro Hospitalar Universitário de Marrakech.

2.4 Grafos de Conhecimento (GC)

De acordo com Pan et al. (2017) um GC possui restrições em sua estrutura impostas por meio de ontologias. Possibilitam a inferência de novos fatos enriquecendo o conhecimento e o seu compartilhamento. Os GCs são frequentemente constituídos a partir de várias fontes de dados (*datasets*) dependendo das técnicas que permitem extrair e ingerir as fontes de dados para geração dos grafos. O resultado é a diversificação de termos e relações entre os dados. A manipulação da diversidade de dados utiliza as representações de esquema que contém as identificações e o contexto das entidades presentes no GC. Esse esquema define uma estrutura de alto nível, onde a identidade dos objetos denota quais nós no grafo se referem à mesma entidade do mundo real, enquanto o contexto especifica os objetos

considerados verdadeiros para o domínio específico. Também são considerados como grafos de dados, destinados a acumular e transmitir conhecimento do mundo real, cujos nós representam entidades de interesse e cujas bordas representam relações entre essas entidades (HOGAN et al., 2020). Hogan et al. (2020) destaca a importância das ontologias como modelos conceituais utilizados para anotar conjuntos de dados. Esta anotação permite gerar fragmentos de conhecimento do domínio que podem ser representados por GCs. Bax e Silva (2020) abordam o uso de ontologias para anotação de dados com vistas a harmonizar e normalizar variáveis de estudos científicos. Propõem que os dados estejam organizados em grafos RDF para serem explorados por consultas a fim de darem origem a outros arquivos de dados.

3 ANOTANDO DADOS COM DICIONÁRIOS SEMÂNTICOS DE DADOS

O SDD é um conjunto de padrões de metadados fundamentados em ontologias que descrevem objetos (representados por dados) em classes e relacionamentos (RASHID et. al, 2017). A anotação por SDD é feita de forma manual e associa os dados de um *dataset* a conceitos (ou classes) nas ontologias, formalizando a semântica dos dados. A anotação dos dados pode ser realizada por profissionais familiarizados com o domínio e os *datasets*, com apoio de ontologistas. A formalização do vocabulário abre caminho para a interoperabilidade dos dados que podem ser integrados de fontes diversas. Abaixo estão relacionados os artefatos utilizados na anotação por SDD:

- Ontologia de domínio. Formaliza os conceitos do domínio da pesquisa. Deve-se buscar reutilizar ontologias consolidadas.
- Dictionary Mapping (DM). Anota as colunas do *dataset*. Cada linha do DM mapeia uma coluna do *dataset*, formalizando-a conceitualmente, explicitando suas relações com os outros dados, bem como a sua proveniência.
- CodeBook. Estrutura os dados categóricos de um *dataset*, mapeando-os para conceitos correspondentes na ontologia.
- Infosheet. Organiza os metadados de descrição do SDD.
- GC (RDF). Resulta da interpretação da dupla: "SDD (*templates* de metadados) + Dados" pelo *script* de anotação. Caso seja necessário persistir os dados, o usuário pode armazenar o grafo em um *triplestore* para consulta posterior dos dados.

Inicialmente, os dados mapeados para as ontologias pelo SDD são as colunas do próprio *dataset*. Os objetos caracterizados nos *datasets* podem estar implicitamente representados. Os objetos implícitos serão explicitados no SDD e formalizados no grafo final gerado. A explicitação dos objetos implícitos favorece a integração semântica dos dados nos níveis conceituais mais abstratos do domínio do projeto, permitindo normalizar e harmonizar interpretações de conceitos que descrevem dados que se deseja integrar.

3.2 Tecnologia para ingestão de dados com base no SDD

Uma vez que a estrutura de anotação esteja pronta, a ferramenta *sdd2rdf* é utilizada para integrar os dados do *dataset* descrito pelo SDD formando um GC persistido em *triplestores*. A *sdd2rdf*⁶ é um *script/software* que interpreta o SDD e converte os dados do *dataset* descrito pelo SDD gerando o GC expresso no padrão RDF (RASHID et. al, 2017). Para exemplificar o acesso aos dados anotados no grafo, o *sdd2rdf* cria alguns exemplos de consultas SPARQL⁷. O GC possibilita a inferência de novos fatos enriquecendo o compartilhamento do conhecimento. Consultas em SPARQL e regras SWRL⁸ permitem inferências. O grafo RDF gerado pelo se fundamenta em ontologias, e possibilita a integração semântica dos dados bem como a sua interoperabilidade.

4 USO DE DICIONÁRIO SEMÂNTICOS DE DADOS NA ESTRATIFICAÇÃO DE RISCOS EM SAÚDE MENTAL

O uso de SDDs para Estratificação de Riscos em Saúde Mental⁹ é resultado da colaboração entre o Hospital Vila Verde Saúde Mental¹⁰ (HVVSM) situado em Juiz de Fora-MG, com o PPGGOC/UFMG¹¹. Essa colaboração permitiu o registro e a realização de um projeto financiado pela FAPEMIG e teve a duração de 9 meses. A equipe multiprofissional composta por um médico psiquiatra, uma enfermeira e uma psicóloga do HVVSM, utiliza um questionário para estratificar os riscos em saúde mental, estruturado por questões organizadas em grupos e organizados por sintomas relacionados aos transtornos mentais. Aplica-se a estratificação de riscos a partir dos dados pessoais, clínicos e aspectos psicossociais. O paciente responde “Sim” ou “Não”. Respostas “Sim” recebem uma

⁶ <https://github.com/tetherless-world/SemanticDataDictionary>

⁷ SPARQL *Protocol and RDF Query Language*

⁸ *Semantic Web Rule Language*

⁹ <https://github.com/evaldo/estratificacaoderiscosaudemental/wiki>

¹⁰ Disponível em: www.vilaverdejf.com.br

¹¹ Programa de Pós-Graduação em Gestão e Organização do Conhecimento da Universidade Federal de Minas Gerais. Disponível em: ppggoc.eci.ufmg.br

pontuação, 0 (zero) para caso contrário. Ao final, o somatório é atribuído a faixa de risco: “Baixo”, “Médio” e “Alto”. A mensuração da presença e gravidade dos sinais e sintomas auxilia o planejamento da assistência em saúde e gera o Projeto Terapêutico Singular. *Datasets* com dados fictícios oriundos da estratificação serão anotados neste artigo (Tabela 1).

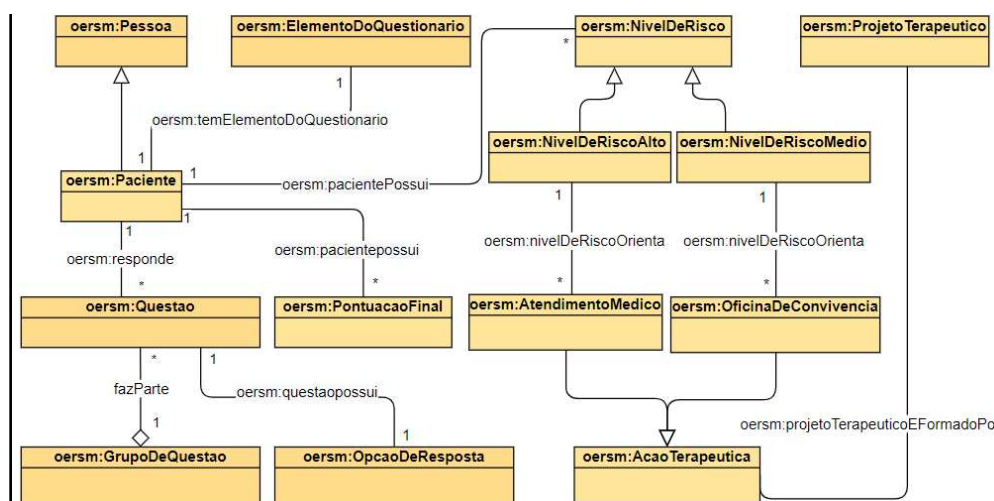
Tabela 1: Dataset com dados da estratificação.

ID do Paciente	Nome do Paciente	Grupo1_1_Sensacao de morte iminente, panico e/ou medo intenso	pontos_grupo1_1	Total	Risco
4	A_25	Nao	15	30	BAIXO
6	A_161	Sim	58	85	ALTO
12	A_130	Nao	32	63	MEDIO

Fonte: Elaborado pelo autor.

A anotação é feita com o uso da Ontologia da Estratificação de Riscos em Saúde Mental (OERSM) (Figura 2). A OERSM foi modelada a partir do entendimento do domínio em reuniões com a equipe assistencial do hospital, e com base no referencial teórico.

Figura 2: Fragmento da OERSM.



Fonte: Elaborado pelo autor.

As colunas da tabela (Tabela 1) são mapeadas para conceitos da OERSM. Questões de competência ilustram problemas que poderiam ser respondidos pela análise das variáveis na Tabela 1. (i) Quais questões foram respondidas por grupo? (ii) Quais planos de cuidados em saúde mental estão relacionados aos pacientes? Os templates de metadados em Rashid et al. (2017) incluem o *Dictionary Mapping* (DM), elaborado nas Tabelas 2 e 3. Os conceitos utilizados para anotação do cabeçalho das tabelas foram reutilizados da ontologia OERSM, formando a ontologia base deste estudo específico. A Tabela 4 apresenta o *CodeBook*.

Tabela 2: Especificação do DM para dados explícitos e implícitos.

Column	Attribute	attributeOf	Unit
ID do Paciente	oersm:IdentificadorDeCadastro	??paciente	schema:Interger
Nome do Paciente	oersm:NomeCompleto	??paciente	
Grupo1_1_Sensacao de morte iminente, panico e/ou medo intenso	oersm:OpcaoDeResposta	??questao	

pontos_grupo1_1	oersm:Pontuacao	??questao	schema:Interger
Total	oersm:PontuacaoFinal	??elementoquestionario	schema:Interger
Risco	oersm:NivelDeRisco	??paciente	

Fonte: Elaborado pelo autor.

Tabela 3: Especificação do DM para explicitação dos dados.

Column	Entity	Relation	inRelationTo
??questao	oersm:Questao	oersm:questaoPossui	oersm:OpcaoDeResposta
??paciente	oersm:Paciente	oersm:responde	oersm:Questao
??paciente	oersm:Paciente	oersm:pacientePossui	oersm:PontuacaoFinal
??elementoquestionario	oersm:ElementoQuestionario	oersm:temElementoQuestionario	oersm:Paciente
??paciente	oersm:Paciente	oersm:pacientePossui	oersm:NivelDeRisco
??nivelderisco	oersm:NivelDeRisco	oersm:nivelDeRiscoOrienta	oersm:ProjetoTerapeutico
??projetoterapeutico	oersm:ProjetoTerapeutico	oersm:projetoTerapeuticoEFormadoPor	oersm:AcaoTerapeutica
??nivelderiscoalto	oersm:NivelDeRiscoAlto	oersm:nivelDeRiscoOrienta	oersm:AtendimentoMedico
??nivelderiscomedio	oersm:NivelDeRiscoMedio	oersm:nivelDeRiscoOrienta	oersm:OficinaDeConvivencia
??nivelderiscobaixo	oersm:NivelDeRiscoBaixo	oersm:nivelDeRiscoOrienta	oersm:InstrumentoAssistencial

Fonte: Elaborado pelo autor.

Tabela 4. Codebook.

Column	Code	Label	Class
RISCO	1	Baixo	:Baixo
RISCO	2	Medio	:Medio
RISCO	3	Alto	:Alto

Fonte: Elaborado pelo autor.

A Figura 03 apresenta o grafo RDF, em sintaxe *turtle*, gerado pelo script *sdd2rdf*, tendo como entradas os metadados das Tabelas 2 e 3 e a primeira linha dos arquivos de dados da Tabela 1.

Figura 3: Grafo RDF no formato *turtle*.

@prefix np: <http://www.nanopub.org/nschema#> .	oersm:paciente a oersm:Paciente,
@prefix ns1: <http://purl.org/dc/terms/> .	oersm:paciente ;
@prefix oersm: <http://hadatac.org/ont/oersm#> .	rdfs:label "paciente"^^xsd:string ;
@prefix owl: <http://www.w3.org/2002/07/owl#> .	oersm:pacientePossui oersm:PontuacaoFinal .
@prefix prov: <http://www.w3.org/ns/prov#> .	oersm:pontos_grupo1_1 a oersm:Pontuacao,
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .	oersm:pontos_grupo1_1 ;
@prefix sio: <http://semanticscience.org/resource/> .	sio:hasUnit sio:Interger ;
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .	sio:hasValue 30 ;
oersm:Grupo1_1_Sensacao_de_morte_iminente__panico_e-ou_medo_intenso a owl:Class ;	sio:isAttributeOf oersm:questao .
rdfs:subClassOf [a owl:Restriction ;	oersm:projetoterapeuticosingular a owl:Class ;
owl:allValuesFrom oersm:questao ;	rdfs:label "projetoterapeuticosingular"^^xsd:string ;
owl:onProperty sio:isAttributeOf],	oersm:projetoTerapeuticoSingularEFormadoPor
oersm:Questao .	oersm:AcaoTerapeutica ;
oersm:ID_do_Paciente a oersm:ID_do_Paciente,	rdfs:subClassOf oersm:ProjetoTerapeuticoSingular .
oersm:IdentificadorDeCadastro ;	oersm:questao a oersm:Questao,
sio:hasUnit sio:Interger ;	oersm:questao ;
sio:hasValue 4 ;	rdfs:label "questao"^^xsd:string ;
sio:isAttributeOf oersm:paciente .	oersm:fazParte oersm:GrupoDeQuestao .
oersm:Nome_do_Paciente a oersm:NomeCompleto,	oersm:questao a oersm:Questao,
oersm:Nome_do_Paciente ;	oersm:questao ;
sio:hasValue "A_25"^^xsd:string ;	rdfs:label "questao"^^xsd:string ;
sio:isAttributeOf oersm:paciente .	oersm:questaoPossui oersm:OpcaoDeResposta .
oersm:RISCO a oersm:NivelDeRisco,	oersm:questao a oersm:Questao,
oersm:RISCO ;	oersm:questao ;
sio:isAttributeOf oersm:paciente .	rdfs:label "questao"^^xsd:string ;
oersm:TOTAL a oersm:PontuacaoFinal,	oersm:questaoPossui oersm:Pontuacao .
oersm:TOTAL ;	oersm:elementoquestionario a owl:Class ;
sio:hasUnit sio:Interger ;	rdfs:label "elementoquestionario"^^xsd:string ;
sio:hasValue "BAIXO"^^xsd:string ;	oersm:temElementoQuestionario oersm:Paciente ;
	rdfs:subClassOf oersm:ElementoQuestionario .

<pre> sio:isAttributeOf oersm:elementoquestionario . oersm:nivelderisco a owl:Class ; rdfs:label "nivelderisco"^^xsd:string ; oersm:nivelDeRiscoOrienta oersm:ProjetoTerapeuticoSingular ; rdfs:subClassOf oersm:NivelDeRisco . oersm:nivelderiscoalto a owl:Class ; rdfs:label "nivelderiscoalto"^^xsd:string ; oersm:nivelDeRiscoOrienta oersm:AtendimentoMedico ; rdfs:subClassOf oersm:NivelDeRiscoAlto . oersm:nivelderiscomedio a owl:Class ; rdfs:label "nivelderiscomedio"^^xsd:string ; oersm:nivelDeRiscoOrienta oersm:AtendimentoMedico ; rdfs:subClassOf oersm:NivelDeRiscoMedio . oersm:paciente a oersm:Paciente, oersm:paciente ; rdfs:label "paciente"^^xsd:string ; oersm:responde oersm:Questao . </pre>	<pre> oersm:paciente a oersm:Paciente, oersm:paciente ; rdfs:label "paciente"^^xsd:string ; oersm:pacientePossui oersm:NivelDeRisco . oersm:questao a owl:Class ; rdfs:label "questao"^^xsd:string ; oersm:fazParte oersm:GrupoDeQuestao ; oersm:questaoPossui oersm:OpcaoDeResposta, oersm:Pontuacao ; rdfs:subClassOf oersm:Questao . oersm:paciente a owl:Class ; rdfs:label "paciente"^^xsd:string ; oersm:pacientePossui oersm:NivelDeRisco, oersm:PontuacaoFinal ; oersm:responde oersm:Questao ; rdfs:subClassOf oersm:Paciente . </pre>
--	---

Fonte: Elaborado pelo autor.

Como visto, a preparação de dados, com SDD, no âmbito da estratificação de riscos em saúde mental permite anotar com metadados produzindo GCs que representam os conceitos e as relações entre eles. Os elementos terminais dos grafos são valores que instanciam propriedades de objetos que, por sua vez, populam a estratificação de riscos e são usados para responder às questões de competência. Os arquivos de dados organizados nos grafos contêm os valores em células tabulares e os templates de metadados são compostos de anotações formais de propriedades de objetos. A OERSM formaliza a semântica exigindo que as definições sejam declaradas usando termos consensuados no domínio. É possível navegar pelo GC a fim de responder às questões de competência. A questão (i) pode ser respondida a partir do objeto `??paciente` que é explicitado no conceito `oersm:Paciente`, e que se relaciona com a classe `oersm:Questao` por meio da propriedade `oersm:responde`. A questão (ii) pode ser respondida a partir do objeto `??paciente`, e que se relaciona com a classe `oersm:NivelDeRisco` por meio da propriedade `oersm:pacientePossui`. A classe `oersm:NivelDeRisco` se relaciona com a classe `oersm:ProjetoTerapeutico` por meio da propriedade `oersm:nivelDeRiscoOrienta`. O objeto `??projetoterapeutico` é explicitado pela classe `oersm:ProjetoTerapeutico`, e que permite a recuperação do plano de cuidados a partir da propriedade `oersm:projetoTerapeuticoEFormadoPor`, que relaciona a classe `oersm:ProjetoTerapeutico` com a classe `oersm:AcaoTerapeutica`.

5 TRABALHOS CORRELATOS

Cardoso et al (2021) aplicaram a ontologia Onto-PaRON para anotar os prontuários de 928 pacientes. As anotações semânticas foram transformadas em dados quantitativos. Constataram que a demanda de assistência em saúde depende da forma inicial da doença, do estado motor e da idade do paciente. O estudo permitiu a gestão do cuidado, relacionada ao

estado motor e cognitivo do paciente. Liu et al (2021) propõem um *knowledge graph* para descobrir as associações potenciais entre a microbiota intestinal, neurotransmissores e transtornos mentais. Os autores integraram ontologias na área biomédica por meio do GC MiKG (*Microbiota Knowledge Graph*). Hadzic e Dillon (2009) desenvolveram um modelo ontológico de saúde mental baseado em evidências. Os autores descrevem como a evolução da ontologia pode ser realizada usando um sistema multiagente em combinação com algoritmos de mineração de dados, que sustentam vários experimentos de pesquisa de natureza colaborativa, necessários para os tratamentos de doenças mentais. Yamada et al (2020) propõem uma ontologia para servir de conhecimento comum entre os envolvidos no domínio da saúde mental, para fazer inferências sobre tratamentos, sintomas, métodos de diagnóstico e prevenção, auxiliando os profissionais de saúde nas decisões clínicas. Bonet-Dalmau et al. (2016) descreveram condições crônicas de Saúde Mental a partir de um instrumento de estratificação de risco para a Atenção Integrada em Saúde Mental na Catalunha. Os autores desenvolveram 3 Grupos de Risco para as condições de Saúde Mental. Testaram a consistência e a previsão das necessidades das pessoas identificadas, atendendo às necessidades sociais e envolvendo a rede de cuidados na gestão de planos terapêuticos.

A técnica de anotação semântica para estratificação de risco em saúde mental apresentada neste artigo se relaciona com as diferentes abordagens citadas. Todas utilizam ontologias para fornecer semântica formal aos *datasets*. Ontologias facilitam a normalização (de escalas e unidades de medida) de alguns dados oriundos de estudos realizados independentemente uns dos outros, permitindo a sua integração. O uso do SDD é, contudo, uma vantagem da nossa abordagem, pois ele permite que especialistas de domínio possam anotar os dados eles próprios, manualmente, interagindo com os dados produzidos pela estratificação de riscos por meio de templates de metadados enriquecidos semanticamente pelas ontologias. Este processo permite que os dados sejam mapeados e "ingeridos" em um repositório, organizando-os em GCs no formato RDF. Para fazer o mapeamento o especialista desenvolve a modelagem conceitual, contribuindo para validar a estrutura da ontologia, refinando o conhecimento a cada anotação de dados. Dados de outros *datasets* podem ser integrados. Por exemplo, é possível validar a eficácia do PTS integrando os dados de diferentes evoluções. Esta abordagem abre caminho para o uso da anotação semântica de dados em estudos científicos. O GC gerado é utilizado para expor aos usuários interfaces com filtros sofisticados para extrair novos *datasets* à demanda, podendo também serem usados para

inferir novos conhecimentos e correlações entre sinais, sintomas, doenças mentais, níveis de riscos estratificados com os PTS gerados para os pacientes.

6 CONCLUSÃO

O artigo descreve como anotar dados de estratificações de riscos em saúde mental com o uso de dicionários semânticos, potencialmente reduzindo o alto custo decorrente dos desafios técnicos enfrentados para reutilizar e integrar dados na área da saúde. Como exemplo, utiliza-se o SDD em conjunto com a ontologia OERSM para interpretar os significados dos conceitos a fim de desambiguar e preparar os dados para análises futuras sobre sinais, sintomas e a gravidade dos pacientes. Os dados são organizados em grafos RDF que podem ser explorados por consultas a fim de darem origem a outros *datasets* ou arquivos de dados. Torna-se possível integrar fontes de dados diversas para enriquecer conjuntos de dados que se quer analisar. Com o uso do SDD a anotação semântica pode ser realizada pelos próprios especialistas de domínio, sem a necessidade do uso de linguagens de programação. A proposta do SDD é voltada, no ciclo de vida do *data science*, para as fases iniciais, permitindo a organização conceitual, na análise e preparação de dados em ciência de dados. A partir do nível de risco estratificado, a equipe multiprofissional elabora o PTS para o paciente. Os campos de digitação livre da evolução dos PTS não foram explorados neste trabalho. As evoluções sobre a eficácia dos PTS podem ser descritas usando digitação livre.

Como trabalhos futuros há a necessidade de fazer a anotação semântica dos dados sobre PTS da forma como são registrados, em texto. Os conceitos presentes nestas evoluções podem ser anotados por técnicas de processamento de linguagem natural (NLP) a fim de enriquecer os conjuntos de dados sobre a estratificação de riscos. O esforço para aplicar a NLP justifica-se pela necessidade de identificar classes de dados para validar a qualidade dos PTS em relação ao risco estratificado. As técnicas de NLP extrairão entidades salientes que representem variáveis para além daquelas presentes no *dataset* sobre a estratificação de riscos em saúde mental, tendo como base tarefas de NLP, tais como: *semantic role labelling*, *word sense disambiguation*, *entity linking* e outras.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG), Processo: APQ-03207-16.

REFERÊNCIAS

AMERICAN PSYCHIATRIC ASSOCIATION, A. P. et al. **Diagnostic and statistical manual of mental disorders (DSM-IV)**. Washington: Washington, DC: American psychiatric association, 1994.

BAX, M. P. ; SILVA, Evaldo de Oliveira . **Dicionários Semânticos de Dados para Integrar Dados de Prontuários Eletrônicos de Pacientes (*Semantic Data Dictionary to Integrate Data of Electronic Medical Records*)**. In: Ontobras, 2020, Vitória. Proceedings of the XIII Seminar on Ontology Research

in Brazil and IV Doctoral and Masters Consortium on Ontologies (ONTOBRAS 2020). Aachen: CEUR, 2020. v. 2728. p. 298-303.

BENFARES, Chaymae; EL IDRISSEI, Younès El Bouzekri; HAMID, Karim. **Personalized healthcare system based on ontologies**. In: International Conference on Advanced Intelligent Systems for Sustainable Development. Springer, Cham, 2018. p. 185-196.

BONET-DALMAU, Pere et al. **Risk Stratification as a tool to the Integrated Care in Mental Health in Catalunya**. International Journal of Integrated Care (IJIC), v. 16, n. 6, 2016.

BRASIL. Ministério da Saúde. **Portaria Nº 3.088, de 23 de dezembro de 2011**. Institui a Rede de Atenção Psicossocial para pessoas com sofrimento ou transtorno mental e com necessidades decorrentes do uso de crack, álcool e outras drogas, no âmbito do Sistema Único de Saúde (SUS). Disponível em: http://bvsms.saude.gov.br/bvs/saudelegis/gm/2011/prt3088_23_12_2011_rep.html. Acesso em 14 de abr de 2021.

BRASIL. Ministério da Saúde. Secretaria de Atenção à Saúde. Departamento de Atenção Básica. Saúde mental / Ministério da Saúde, Secretaria de Atenção à Saúde, Departamento de Atenção Básica, Departamento de Ações Programáticas Estratégicas. – Brasília : Ministério da Saúde. **Caderno de Atenção Básica**, 2013.

BRASIL. Presidência da República. Casa Civil. **TelePSI: projeto é ampliado para profissionais de serviços essenciais**. 2020. Disponível em: <https://tinyurl.com/ktebwnc>. Acesso em 10 de abr de 2020.

BRENAS, Jon Hael; SHIN, Eun Kyong; SHABAN-NEJAD, Arash. **Adverse childhood experiences ontology for mental health surveillance, research, and evaluation: advanced knowledge representation and semantic web techniques**. JMIR mental health, v. 6, n. 5, p. e13498, 2019.

CARDOSO, Sonia et al. **Use of a modular ontology and a semantic annotation tool to describe the care pathway of patients with amyotrophic lateral sclerosis in a coordination network**. Plos one, v. 16, n. 1, p. e0244604, 2021.

CEUSTERS, Werner; SMITH, Barry. **Foundations for a realist ontology of mental disease**. Journal of biomedical semantics, v. 1, n. 1, p. 1-23, 2010.

ESPÍRITO SANTO (Estado). **Protocolo de Classificação de Risco em Saúde Mental. Secretaria do Estado do Espírito Santo**. 2018. Disponível em: <https://tinyurl.com/28zxkdbd>. Acesso em 09 de abril de 2021.

HADZIC, Maja; DILLON, Darshan. **An agent-based data mining system for ontology evolution**. In: OTM Confederated International Conferences" On the Move to Meaningful Internet Systems". Springer, Berlin, Heidelberg, 2009. p. 836-847.

HASTINGS, Janna et al. **Representing mental functioning: Ontologies for mental health and disease**. 2012.

HOGAN, Aidan, BLOMQVIST, Eva, COCHEZ, Michael, D'AMATO, Claudia, MELO Gerard de, GUTIERREZ, Claudio, GAYO, José Emilio Labra, KIRANE, Sabrina, NEUMAIER, Sebastian, POLLERES, Axel, NAVIGLI, Roberto, NGOMO, Axel-Cyrille Ngonga, RASHID, Sabbir M., RULA, Anisa, SCHMELZEISEN, Lukas, SEQUEDA, Juan, STAAB, Steffen, ZIMMERMANN, Antoine. **Knowledge Graphs**. arXiv preprint arXiv:2003.02320, 2020.

LIU, Ting et al. **Exploring the Microbiota-Gut-Brain Axis for Mental Disorders with Knowledge Graphs**. Journal of Artificial Intelligence for Medical Sciences, 2020.

MELLO, Veronica de Pádua. **Caminhos da educação em saúde na atenção básica: proposta de reorganização do grupo do parque**. 2015. Tese de Doutorado. Universidade de São Paulo.

MOTTA, Luis; BLAYA, Carolina; COSTA, Marianna De Abreu; PASSOS, Ives. **Telepsiquiatria: Estratificação de Riscos, Transição para o Cuidado Presencial e Telessuporte ao Profissional de Referência - Projeto TELEPSI**. Disponível em: https://www.researchgate.net/publication/344654886_Telepsiquiatria_Estratificacao_de_Riscos_Transicao_para_o_Cuidado_Presencial_e_Tele_suporte_ao_Profissional_de_Referencia_-_Projeto_TELEPSI. Acesso em 09 de abr de 2021.

MILLER, Charles C.; REARDON, Michael J.; SAFI, Hazim J. **Risk stratification: a practical guide for clinicians**. Cambridge University Press, 2001.

OMS. Organização Mundial da Saúde. Organização Mundial de Médicos de Família. **Integração da saúde mental nos cuidados de saúde primários: uma perspectiva global**. 2008. Disponível em: https://www.who.int/eportuguese/publications/Integracao_saude_mental_cuidados_primarios.pdf. Acesso em : 17 de abr de 2021.

PAN, Jeff Z. et al. (Ed.). **Exploiting linked data and knowledge graphs in large organisations**. Heidelberg: Springer, 2017.

PARANÁ (Estado). **Rede de Saúde Mental**. Disponível em: <https://www.saude.pr.gov.br/Pagina/Saude-Mental> Acesso em 09 de abr de 2021.

PAULA, George Luiz Costa de. **Classificação de risco em saúde mental: implicações clínicas, éticas e sócio-políticas**. 2019.

MICROSOFT. **O estágio de entendimento empresarial do ciclo de vida do Processo de Ciência de Dados da Equipe**. 2020. Disponível em: <https://docs.microsoft.com/pt-br/azure/machine-learning/team-data-science-process/lifecycle-business-understanding>. Acesso em: 17 de maio de 2021.

NCISH, S, SiM H. **The assessment of clinical risk in mental health services. Manchester: The University of Manchester**, 2018. Disponível em: <https://documents.manchester.ac.uk/display.aspx?DocID=38466>. Acesso em 09 de abr de 2021.

RASHID, Sabbir M. et al. **The Semantic Data Dictionary Approach to Data Annotation & Integration**. In: SemSci@ ISWC. 2017. p. 47-54.

RASHID, S. M., MCCUSKER, J. P., PINHEIRO, P., BAX, M. P., SANTOS, H., STINGONE, J. A., ... & MCGUINNESS, D. L. (2020). **The Semantic Data Dictionary—An Approach for Describing and Annotating Data**. Data Intelligence, 443-486.

SAXENA, Shekhar; SETOYA, Yutaro. **World Health Organization's Comprehensive Mental Health Action Plan 2013–2020**. 2014.

YAMADA, Diego Bettiol et al. **Ontology-Based Inference for Supporting Clinical Decisions in Mental Health**. In: International Conference on Computational Science. Springer, Cham, 2020. p. 363-375.

WHO. World Health Organization Geneva. **Classificação de transtornos mentais e de comportamento da CID - 10: descrições clínicas e diretrizes diagnósticas**. Porto Alegre: Artmed, 1993.