

UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE CIÊNCIAS ECONÔMICAS
CENTRO DE DESENVOLVIMENTO E PLANEJAMENTO REGIONAL

SAHRA FERREIRA PINHEIRO

***A WORLD WIDE WEB COMO UMA GENERAL PURPOSE TECHNOLOGY: UMA
PROPOSTA DE INVESTIGAÇÃO.***

**Belo Horizonte
2022**

Sahra Ferreira Pinheiro

A World Wide Web como uma General Purpose Technology:
uma proposta de investigação.

Tese apresentada ao Centro de Desenvolvimento e Planejamento Urbano da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Doutor em Economia.

Orientador(a): Profa. Dra. Márcia Siqueira Rapini.

Co-orientador: Prof. Dr. Leonardo Costa Ribeiro.

Belo Horizonte
2022

Ficha catalográfica

P654w
2022 Pinheiro, Sahra Ferreira.
A World Wide Web como general purpose technology
[manuscrito]: uma proposta de investigação / Sahra Ferreira Pinheiro.
– 2022.
143 f.

Orientadora: Márcia Siqueira Rapini.

Coorientador: Leonardo Costa Ribeiro.

Tese (doutorado) – Universidade Federal de Minas Gerais, Centro
de Desenvolvimento e Planejamento Regional.

Inclui bibliografia (fls. 138-143).

1. World Wide Web – servidores – Teses. 2. Redes de informação
– Teses. 3. Tecnologia e desenvolvimento econômico – Teses. I.
Rapini, Márcia Siqueira. II. Ribeiro, Leonardo Costa. III.
Universidade Federal de Minas Gerais. Centro de Desenvolvimento e
Planejamento Regional. IV. Título.

CDD: 303.483

Elaborado por Rosilene Santos CRB-6/2527
Biblioteca da FACE/UFMG. –RSS65/2023



UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE CIÊNCIAS ECONÔMICAS
CENTRO DE DESENVOLVIMENTO E PLANEJAMENTO REGIONAL
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

FOLHA DE APROVAÇÃO

SAHRA FERREIRA PINHEIRO

A WORLD WIDE WEB COMO UMA GENERAL PURPOSE TECHNOLOGY: UMA PROPOSTA DE INVESTIGAÇÃO

Tese apresentada ao Programa de Pós-Graduação em Economia, da Faculdade de Ciências Econômicas da Universidade Federal de Minas Gerais, para obtenção do título de Doutor em Economia, área de concentração em Economia.

Aprovada em Belo Horizonte, 07 de julho de 2022.

BANCA EXAMINADORA – participações por videoconferência:

Profa. Márcia Siqueira Rapini (Orientadora) (Cedeplar/FACE/UFMG)

Prof. Leonardo Costa Ribeiro (Coorientador) (Cedeplar/FACE/UFMG)

Prof. Eduardo da Motta e Albuquerque (Cedeplar/FACE/UFMG)

Profa. Ana Lúcia Tatsch (UFRGS)

Prof. Gilberto Medeiros Ribeiro (DCC/UFMG)

Prof. Túlio Chiarini (IPEA)

ÉDSON PAULO DOMINGUES
Coordenador do Programa de Pós-Graduação em Economia



Documento assinado eletronicamente por **Marcia Siqueira Rapini, Professora do Magistério Superior**, em 07/07/2022, às 17:35, 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 **Leonardo Costa Ribeiro, Professor do Magistério Superior**, em 08/07/2022, às 14:54, 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 **Gilberto Medeiros Ribeiro, Membro de comissão**, em 12/07/2022, às 10:32, 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 **Eduardo da Motta e Albuquerque, Professor do Magistério Superior**, em 13/07/2022, às 11:25, 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 Lúcia Tatsch, Usuário Externo**, em 13/07/2022, às 12:06, 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 **Túlio Chiarini de Faria, Usuário Externo**, em 13/07/2022, às 12:38, 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 **Edson Paulo Domingues, Coordenador(a) de curso de pós-graduação**, em 13/07/2022, às 16:03, 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 **1562772** e o código CRC **278A8670**.

RESUMO

O presente trabalho tem por objetivo investigar se o surgimento da *World Wide Web* conformou um sistema complexo de tecnologia com características próprias de uma tecnologia do tipo *General Purpose Tecnnology*. A hipótese analisada é a de que o conhecimento presente no artigo inaugural da WWW (BERNERS-LEE, 1992) possui relevância tal que causou a explosão de numerosos progressos tecnológicos posteriores e que esse movimento pode ser captado (ou observado) pelas ondas de patenteamento e de publicação científica subsequentes. Para tanto, a partir de dados retirados do PATSTAT e da *Web of Science*, construiu-se duas redes, uma com 1,5 milhões de patentes e outra com 13,4 milhões de artigos, e suas respectivas citações, para captar as dinâmicas tecnológicas e científicas e a identificação das semelhanças com as estruturas clássicas de uma GPT. O período de análise compreende, aproximadamente, 20 anos, com início na data de publicação do artigo de divulgação da WWW de Berners-Lee em 1992, até 2012. A partir das ferramentas da “ciência de redes”, observam-se redes com distribuição com lei de potência com propriedades *scale-free*, sendo compatível com as previsões da literatura para redes tecnológicas e científicas, isto é, muitos nós com poucas conexões em contrapartida a poucos nós com número elevado de conexões, determinando de sobremaneira a forma e dinâmica de comportamento da rede. Ademais, característica *small world* também foram encontradas, indicando a alta conectividade na rede. Esses resultados sustentam a hipótese proposta e apontam que a WWW possui as características de uma tecnologias do tipo GPT, sendo definidas como tecnologias de caráter geral e alta pervasividade. Estas tecnologias possuem alta capacidade de aplicação nos mais variados setores de atividade, sendo encabeçadas por alguns clusters de tecnologia chaves, e com capacidade de fomentar a inovação nos mais diversos setores.

Palavras-chave: World Wide Web. Paradigmas tecnológicos. Trajetórias Tecnológicas. General Purpose Technology. Ciência de redes.

ABSTRACT

This work aims to investigate whether the emergence of the World Wide Web shaped a complex technology system with characteristics typical of a General Purpose Technology. The hypothesis analyzed is that the knowledge present in the WWW's inaugural article (Berners-Lee, 1992) has such relevance that it caused the explosion of numerous subsequent technological advances and that this movement can be captured by the waves of patenting and subsequent scientific publication. To this end, based on data taken from PATSTAT and the Web of Science, two networks were built, one with 1.5 million patents and the other with 13.4 million articles, and their respective citations, to capture the technological dynamics and scientific and the identification of similarities with the classic structures of a GPT. The period of analysis covers approximately 20 years, starting from the date of publication of Berners-Lee's WWW article in 1992, until 2012. From the tools of "network science", networks with power law distribution with scale-free properties, being compatible with the predictions in the literature for technology and scientific networks, that is, many nodes with few connections in contrast to few nodes with a high number of connections, greatly determining the form and dynamics of network behavior. Furthermore, small world characteristics were also found, indicating high connectivity in the network. These results support the proposed hypothesis and point out that the WWW has the characteristics of GPT technologies, being defined as technologies of a general nature and high pervasiveness. Which have a high capacity for application in the most varied sectors of activity, being headed by some key technology clusters, and capable of fostering innovation in the most diverse sectors.

Keywords: World Wide Web. Technological paradigms. Technological Trajectories. General Purpose Technology. Network science.

SUMÁRIO

INTRODUÇÃO	10
CAPÍTULO 1: PARADIGMAS TECNOLÓGICOS, TRAJETÓRIAS E <i>GENERAL PURPOSE TECHNOLOGIES</i>: UMA PROPOSTA TEÓRICA PARA A LEITURA DA <i>WORLD WIDE WEB</i>.....	14
1.1. Paradigmas e trajetórias tecnológicas.....	16
1.2. General Purpose Technologies.....	29
1.3. Considerações finais.....	37
CAPÍTULO 2: A <i>WORLD WIDE WEB</i>.....	40
2.1. A Internet: contextualizando.....	40
2.2. <i>World Wide Web</i>: antecedentes, criação e desenvolvimentos posteriores.....	47
2.2.1. <i>Antecedentes</i>	48
2.2.2. <i>A invenção: o processo</i>	56
2.2.3. <i>A difusão: o início</i>	60
2.2.4 <i>O surgimento de novas indústrias, setores de atividade e modelos de negócios: o caso da Google</i>	66
2.3. Considerações finais.....	71
CAPÍTULO 3: PROPOSTA METODOLÓGICA PARA ANÁLISE DA TRAJETÓRIA DA <i>WORLD WIDE WEB</i>.....	73
3.1. Patentes e citações de patentes como indicadores: uma revisão conceitual e metodológica.....	73
3.1.1. <i>Da literatura sobre uso de patentes e citações: uma breve síntese</i>	77
3.2. Artigos como indicadores da dinâmica científica: uma revisão conceitual e metodológica.....	85
3.2.1. <i>Revisão de estudo bibliométricos para análise da dinâmica e evolução do conhecimento científico</i>	88
3.3. Da literatura sobre a investigação empírica de <i>General Purpose Technologies</i>.....	94
3.4. As redes de patentes e artigos da <i>World Wide Web</i>.....	96
3.4.1. <i>A rede de patentes</i>	96
3.4.2. <i>A rede de artigos</i>	99

3.5. A análise de redes: <i>Science networks</i>.....	100
3.5.1. <i>Sistemas complexos e ciência de redes</i>	100
3.5.2. <i>Redes: definições</i>	102
3.5.3. <i>Redes aleatórias (random networks)</i>	106
3.5.4. <i>Scale-free networks</i>	107
CAPÍTULO 4: A <i>WORLD WIDE WEB</i> COMO UMA <i>GENERAL PURPOSE TECHNOLOGY</i>.....	111
4.1. Análise descritiva dos dados: investigação preliminar.....	111
4.1.1. <i>Rede de patentes</i>	112
4.1.2. <i>Rede de artigos</i>	121
4.2. Análise das redes: <i>Science Networks</i>.....	124
4.2.1. <i>Rede de patentes da World Wide Web</i>	125
4.2.2. <i>Rede de artigos da WWW</i>	130
4.3. A <i>World Wide Web</i> como uma <i>General Purpose Technology</i>: uma proposta de interpretação.....	133
4.4. Considerações finais.....	136
CONCLUSÃO.....	137
REFERÊNCIAS BIBLIOGRÁFICAS.....	140

INTRODUÇÃO

A fase atual do capitalismo global tem, frequentemente, ganhado denominações como a era da informação, do capitalismo digital, do capitalismo de plataformas ou do *surveillance capitalism* (ALBUQUERQUE, 2021, 2019; SRNICEK, 2017; ZUBOFF, 2019). Essa flagrante dinâmica digital está diretamente relacionada ao surgimento da Internet, mais especificamente à sua popularização via democratização do acesso, bem como de seu uso comercial, permitido pela invenção da *World Wide Web* (WWW) no início da década de 1990 (BERNERS-LEE *et al.*, 1992; BERNERS-LEE E FISCHETTI, 2000; GILLIES E CAILLIAU, 2000; GREENSTEIN, 2015). Inicialmente pensada e criada para promover maior integração entre os cientistas das diversas instituições e universidades (CERN, 2020), potencializando o intercâmbio de conhecimento, essa invenção é resultado de um mundo em transformação, derivada de mudanças técnicas anteriores e em curso, que criaram as condições e a demanda para uma maior conectividade – em especial em um contexto de revolução tecnológica das TICs e do paradigma tecno-econômico vigente da informação e telecomunicações (PEREZ, 2004). A posterior disponibilização do uso livre da rede gerou profundas transformações na vida dos indivíduos e na economia como um todo. De fato, na atualidade, o uso da rede é, a tal ponto, tão corriqueiro e parte integrante do funcionamento da sociedade, em seus mais variados aspectos, que pouco se questiona sobre sua origem e o modo pelo qual sua difusão causou e causa numerosas transformações significativas.

Um dos efeitos dessa transformação é o surgimento de novas mercadorias, como os mecanismos de busca na Internet, dando lugar a criação de novas empresas como a Google, bem como de novas formas de organização do mercado e do consumo, exemplificadas por empresas como a Amazon. Novas formas de se relacionar entre os indivíduos e entre estes e o mercado. Ademais, essa nova tecnologia parece ditar a forma como se inova: novos desenvolvimentos que potencializam o uso da WWW e sendo, por sua, vez potencializados por ela. Ademais, assistimos nas últimas décadas às transformações das instituições para acompanhar o novo cenário imposto pela difusão da conectividade, como as leis de regulação de conteúdo e de acesso à informação. Ou seja, a inovação da WWW parece ter impactos nas esferas econômicas, institucionais e sociais (ALBUQUERQUE, 2021, 2019; SRNICEK, 2017; ZUBOFF, 2019; GREENSTEIN, 2015).

Atualmente, verifica-se a criação de oportunidades para novas invenções e inovações com o surgimento da *big data*, empresas como o Facebook, Google, Amazon etc., tem investido em tecnologias de inteligência artificial (AI) que decorrem do surgimento de um elevado volume de dados com crescimento exponencial e da necessidade de processá-los, criando novas formas de interação entre clientes e empresas, por exemplo. Tecnologias como o *chat box* de sites de prestadores de serviços - como bancos, telecomunicações, entre outros - e outros dispositivos, como a *Bia* do banco Itaú, o dispositivo da Google Home, o Alexa da Amazon etc. Além do mais, termos como Internet das coisas (IoT) e *Blockchains* se tornaram comuns em nossos dias (ALBUQUERQUE, 2021, 2019; GREENSTEIN, 2015; SRNICEK, 2017; ZUBOFF, 2019;).

A percepção de um aparente potencial transformador e indutor da dinâmica inovativa, bem como de seus efeitos e processos retroalimentativos com respeito aos aspectos econômicos, sociais e institucionais, no geral, fornece o objeto de estudo para este trabalho que se debruçará sobre a análise da invenção da WWW e sua subsequente difusão, evocando os desenvolvimentos teóricos sobre paradigmas e trajetórias tecnológicas (DOSI, 1982, 1988), bem como aqueles que tratam das inovações denominadas de *General Purpose Technology* (BRESNAHAN E TRAJTENBERG, 1992, 1995; BEKAR, CARLAW e LIPSEY, 2018). Segundo Albuquerque (2019), a invenção da *World Wide Web* (WWW) sinaliza um evento decisivo para o surgimento de uma nova fase do capitalismo global, podendo configurar uma revolução tecnológica, dado que possui impacto similar ao de outras inovações radicais desencadeadoras de *big bangs* na dinâmica tecnológica, a saber: a estrada de ferro, a máquina a vapor, eletricidade e os microprocessadores. No entanto, o presente trabalho se delimita ao objetivo, menos audacioso, de analisar a trajetória tecnológica da Web, desde sua invenção, a partir de seus aspectos tecnológicos e científicos, buscando identificar seus efeitos e alcance, ou seja, buscando delimitar nesta dinâmica a existência de uma tecnologia de caráter geral. Dito de outra forma, a hipótese proposta é a de que a invenção da WWW inaugura uma trajetória tecnológica própria e relevante que possui efeitos sobre os demais setores da economia, na esfera social e institucional, dado seu caráter de uso geral.

A partir dessas considerações, o presente trabalho se propõe a examinar se a invenção da WWW em 1991, formalizada através do artigo publicado por Berners-Lee (1992), conformou um sistema complexo de tecnologia com características própria de uma *general purpose technology* (GPT). Para tanto, propõe-se a investigação da trajetória de uma tecnologia em específico, através da análise do padrão de formação das redes de patentes e artigos, na busca de identificar

semelhanças com as estruturas clássicas das GPTs. A hipótese subjacente é a de que um conhecimento tecnológico, disponível em um artigo, possui relevância tal que causou a explosão de numerosos progressos tecnológicos posteriores e que esse movimento poderia ser captado pela onda de patenteamento e publicações científicas subsequentes. Para a identificação de uma GPT busca-se explorar as relações entre esta tecnologia e o surgimento de seus setores de aplicação ou as tecnologias que foram possibilitadas através dela, bem como seu potencial de transformação nas esferas econômicas, sociais e institucionais. Segundo Albuquerque (2019), existe uma demanda pela exploração das características potenciais de uma nova revolução tecnológica aparentemente em curso, baseada nas tecnologias de informação e comunicação. Ademais, nota-se a emergência de um conjunto de novas tecnologias candidatas a GPTs, com inúmeras possibilidades de combinação entre elas, abrindo a possibilidade de discussão sobre o impacto potencial dessas tecnologias.

O uso de redes para representação de sistemas complexos, em especial os de tecnologia, foi amplamente discutido por Barabási (2016), portanto, para a investigação definida, o presente trabalho propõe a construção de duas redes: uma contendo 1,5 milhões de patentes sem repetições e um total de citações de patentes na ordem dos 6,3 milhões; a outra somando 13,4 milhões de artigos e um número aproximado de links da rede (ou seja, o número de citações entre artigos) de 109,2 milhões. Ambas as redes possuem como nó inicial o artigo de Berners-Lee de 1992, considerado o artigo que formaliza os princípios básicos contidos na tecnologia da WWW, porém, a rede de patentes contém aquelas que foram concedidas até 2012, cobrindo um período de 20 anos, e a de artigos incluindo publicações até o ano de 2015. Os dados utilizados foram retirados do PATSTAT, no caso das patentes, e da *Web of Science*, para as publicações científicas.

Às redes criadas foram aplicadas as ferramentas da “ciência de rede” (BARABÁSI, 2016; NEWMAN, 2018; MENCZER ET AL., 2020; LATORA ET AL., 2017; CALDARELLI E CHESSA, 2016) para diagnóstico e interpretação de seus respectivos padrões - características e topologia -, resultando na identificação de redes com distribuição lei de potência (*power law*) com propriedades *scale-free*, isto é, muitos nós com poucas conexões em contrapartida a poucos nós com um número elevado de conexões. Esses *hubs*, por sua vez, determinando sobremaneira a forma e a dinâmica de comportamento da rede. Características do tipo *small world* também foram encontradas, indicando a alta conectividade na rede, o que aponta para a potencial presença de uma tecnologia de alta pervasividade. A propriedade *small world* garante uma

rápida e eficiente difusão do conhecimento, tanto tecnológico quanto científico. Ademais, a presença de *hubs* confere à rede outro aspecto importante que trata de sua robustez, isto é, a dificuldade de desmantelamento da rede a partir de retirada de nós aleatórios, ou seja, a difusão do conhecimento tende a permanecer por um período longo de tempo. Esses resultados sustentam a hipótese proposta e apontam que a WWW possui as características de uma GPT.

Por fim, a presente tese encontra-se organizada da seguinte forma, o primeiro capítulo apresenta uma revisão dos desenvolvimentos teóricos acerca dos conceitos de paradigmas tecnológicos, trajetórias tecnológicas e *general purpose technologies*. para o estabelecimento da base teórica sobre a qual se dará a investigação proposta no trabalho e, principalmente, para a delimitação das características próprias destas dinâmicas que nortearão a análise da trajetória da WWW. Em seguida, no segundo capítulo, um breve histórico da invenção e posterior difusão da Web toma lugar, bem como sua conceitualização, com a descrição de seu contexto tecnológico, de seu antecedente e seus desenvolvimentos posteriores. O terceiro capítulo discute metodologicamente o uso de patentes e artigos como indicadores das dinâmicas tecnológicas e científicas, respectivamente, seguindo para a apresentação dos dados e da metodologia proposta de compilação e tratamento destes. Por fim, o quarto capítulo se ocupará da apresentação dos primeiros resultados e de uma proposta interpretativa das mudanças ocorridas com a criação da Web e das que ainda estão em curso sob o efeito desta tecnologia de propósito geral.

CAPÍTULO 1: PARADIGMAS TECNOLÓGICOS, TRAJETÓRIAS E *GENERAL PURPOSE TECHNOLOGIES*.

O conceito de paradigma tecnológico foi introduzido por Dosi (1982) a partir de sua análise da mudança técnica e da percepção de que esta apresenta um padrão recorrente. O paradigma tecnológico, dessa forma, incorpora as diretrizes que guiam o processo evolutivo de mudança técnica através de algumas trajetórias tecnológicas chaves. As oportunidades inovativas se encontram circunscritas dentro do universo de possibilidades que cada paradigma conforma. A inovação, então, segundo esta abordagem, não seria um processo isolado e descolado das tecnologias existentes, mas, antes, um processo cumulativo e inteiramente influenciado pelo contexto econômico, social e institucional em que se encontra.

Entendida como uma atividade solucionadora de problemas, a inovação se baseia no arsenal de conhecimento disponível e selecionado para atingir a solução desejada. As oportunidades inovativas adviriam do paradigma tecnológico em voga, ou seja, do conjunto de princípios básicos governando o avanço tecnológico na firma, na indústria, no setor, na região, ou na época em questão. Tal paradigma informaria a melhor base de conhecimento, as melhores ferramentas e as melhores oportunidades a direcionar as atividades inovativas. Ou seja, a inovação tende a acontecer dentro dos limites do paradigma, que, por sua vez, possui mecanismos de exclusão daquilo que julga não ser compatível com seus princípios e componentes. Assim, além de fornecer componentes materiais para a solução dos problemas, o paradigma também circunscribe um conjunto de problemas relevantes a serem resolvidos e quais as direções possíveis de tais soluções, apresentando, assim, uma heurística do progresso tecnológico.

Dentro desse arcabouço teórico, o progresso tecnológico, ou a evolução tecnológica, ocorre através do que se denominou trajetórias tecnológicas. São estas que, em última instância, irão colocar em curso o processo de desenvolvimento tecnológico dentro do paradigma, lidando com os obstáculos e limites impostos pelos desafios tecnológicos que se apresentam na medida em que se difunde a tecnologia. São também conhecidas como as grandes avenidas de inovação dentro de determinado paradigma. O paradigma, eventualmente, alcança a maturidade e, na medida em que novas demandas e circunstâncias se apresentam, ele gradualmente perde a capacidade de solucionar novos problemas de forma satisfatória, dando lugar a outro paradigma e, conseqüentemente, ao surgimento de novas trajetórias.

Existem paradigmas das mais variadas dimensões e grau de generalidade, e estes se constroem, em geral, em torno de uma tecnologia central ou uma tecnologia capacitadora, ou, ainda, em torno de uma *general purpose technology*, crescendo de forma orgânica a partir desse núcleo. No extremo, tem-se as formulações sobre revoluções tecnológicas associadas ao estabelecimento de um novo paradigma tecno-econômico, definido por Perez (2002, 2004, 2010, 2016), Dosi (1982) e Freeman e Perez (1988), como sendo um conjunto de melhores práticas que regulam o sistema composto pela inovação radical e o cluster que ela gera. Deriva da dinâmica das revoluções tecnológicas que as condições para o surgimento da inovação radical vão sendo criadas ao longo do período de maturação e esgotamento do paradigma anterior e cuja difusão se configura em uma nova revolução tecnológica estabelecendo um novo paradigma. Além do impacto dessa invenção se dar em todas as áreas da vida humana (econômicas, sociais e institucionais) e não apenas em áreas correlatas à expansão da própria rede, com a criação de novas indústrias, novos mercados, novas formas de organização industrial, inovações gerenciais, novas instituições de suporte, novas relações sociais e etc.

O presente trabalho se propõe a examinar a trajetória da *World Wide Web* (WWW), invenção surgida em 1991 e formalizada através do artigo publicado por Berners-Lee (1992), com o objetivo de identificar elementos que corroborem a hipótese de ser esta uma *general purpose technology* (GPT). Albuquerque (2019 e 2021) argumenta que a invenção da WWW se assemelharia ao que é denominado *big bang* de uma revolução tecnológica que, conseqüentemente, iniciou uma nova fase no capitalismo global, estabelecendo um novo paradigma tecno-econômico (PEREZ, 2010). Ainda segundo o autor, existe uma demanda pela exploração das características potenciais de uma nova revolução tecnológica aparentemente em curso, baseada nas tecnologias de “informação e comunicação”. Ademais, nota-se, a partir da invenção e difusão da WWW, a emergência de um conjunto de novas tecnologias candidatas a GPTs, com inúmeras possibilidades de combinação entre elas, abrindo a possibilidade de discussão sobre o impacto potencial dessas tecnologias. No entanto, o objetivo deste trabalho é menos ambicioso, focado na investigação da suposição de que uma determinada tecnologia, incorporando uma combinação de conhecimento tecnológico e científico, disponível em um artigo, se configuraria uma GPT e explorar as relações entre esta tecnologia e o surgimento de seus setores de aplicação, ou as tecnologias que foram possibilitadas através dela, bem como seu potencial de transformação das esferas econômicas, sociais e institucionais. Isso, através da análise de suas dinâmicas tecnológicas e científicas, cuja captura se dará pela proposta de utilização das redes de artigos e patentes que o artigo ensejou.

Para tanto, em primeiro lugar, este capítulo se debruçará sobre os aspectos teóricos que irão embasar a investigação sobre a WWW através da análise de sua trajetória tecnológica e de seus possíveis impactos, na tentativa de identificar padrões e/ou características que a classifiquem como sendo uma GPT. Na primeira seção, será apresentada uma revisão da literatura sobre o arcabouço teórico que trata dos paradigmas tecnológicos e das trajetórias tecnológicas, sendo extensamente baseada nas formulações de Dosi (1982 e 1988). A segunda seção se ocupará da abordagem sobre as *General Purpose Technologies*, proposta, inicialmente, por Bresnahan e Trajtenberg (1992 e 1995)¹, que enriqueceram a elaboração sobre a dinâmica tecnológica de longo prazo, ao contribuir para uma visão mais próxima do caráter turbulento e não-linear do progresso tecnológico (ALBUQUERQUE, 2019).

1.1 Paradigmas e trajetórias tecnológicas.

Dosi (1982) define os conceitos de paradigmas tecnológicos e trajetórias tecnológicas a partir da busca por uma teoria que explicasse a ocorrência da mudança técnica, sua natureza, o que determina sua direção, e, principalmente, que superasse as dificuldades encontradas pelas abordagens denominadas “*demand pull*” e “*technology push*”, levando em conta as complexas interações entre as esferas econômicas, sociais e institucionais na determinação da mudança e do progresso tecnológico. Em especial, sua crítica a estes modelos se concentrava, resumidamente, na descrença sobre o papel indutor preponderante e único da demanda sobre a mudança técnica e o caráter apenas reativo da tecnologia em face as alterações na demanda, por um lado, e na falha dos modelos de *technology push* em reconhecer a importância dos fatores econômicos (DOSI, 1982; 1988), por outro lado. Os conceitos de paradigma e trajetórias tecnológicas foram inspirados na definição, elaborada por Thomas Kuhn², de paradigma científico, ou programa de pesquisa científica, dada a observação de que o processo de geração de conhecimento tecnológico guarda bastante similaridades com o de conhecimento científico, de forma mais ampla. Portanto, através de uma analogia com o conceito de paradigma

1 Já existiam discussões teóricas sobre tecnologias com características do que Bresnahan e Trajtenberg (1992) denominam GPT, como, por exemplo, a distinção entre micro e macro inovações de Mokyr (1990). No entanto, a cunhagem do termo e a sistematização de definições e características na literatura datam a partir do trabalho citado.
2 VON TUNZELMANN, MALERBA, NIGHTINGALE E METCALFE (2008) argumentam que o surgimento de uma filosofia da ciência, representada pelas formulações de Kuhn e Lakatos, evidenciou a ausência de uma análise semelhante para a tecnologia.

científico, cria-se os conceitos de paradigma tecnológico, ou programa de pesquisa tecnológica, e de trajetória tecnológica como análogo ao processo de ciência natural.

Com esses conceitos, Dosi (1982) tenta responder às questões sobre o porquê de determinadas tecnologias emergirem ao invés de outras, se havia regularidades no processo de geração de novas tecnologias e no progresso tecnológico destas e se havia regularidades nas interações entre fatores econômicos, sociais, institucionais e científicos quando de sua influência sobre o processo inovativo. Estas indagações orientaram a construção de seu arcabouço conceitual sobre a mudança técnica. Em seu trabalho de 1988, Giovanni Dosi, utilizando-se de uma abordagem mais microeconômica da mudança técnica, procura investigar a natureza do processo inovativo, a partir de três questões orientadoras: quais as principais características do processo inovativo no geral, quais os fatores que levam ao desenvolvimento de novos produtos e processos de produção e quais os processos que determinam a seleção de determinada inovação e seus efeitos sobre a estrutura industrial. Ou seja, o autor foca em dois grupos de questões: a caracterização do processo inovativo no geral e os fatores que determinam os modos de busca inovativa e as taxas de inovação. Aprofundando, assim, sua elaboração sobre paradigmas e trajetórias tecnológicas. Dosi e Orsenigo (1988) argumentam que apesar da diversidade de fontes e consequências da mudança técnica, este não é um processo puramente aleatório. Existem regularidades no padrão de mudança técnica, apontando a existência de trajetórias tecnológicas e paradigmas tecnológicos, que oferecem oportunidade de investimentos inovadores lucrativos e crescimento de novos mercados ao longo de extensos períodos e caminhos bem definidos de desenvolvimento e difusão.

O primeiro passo de sua elaboração conceitual constitui-se na apresentação de uma definição de tecnologia mais ampla do que a considerada na teoria econômica, principalmente a de inclinação neoclássica, através de uma analogia com os processos científicos. Assim, define tecnologia como sendo “um conjunto de pedaços de conhecimentos, tanto diretamente práticos (relacionados à problemas concretos e dispositivos) e teóricos (porém praticamente aplicáveis apesar de não necessariamente já aplicados), know-how, métodos, procedimentos, experiências de sucesso e fracasso e, também, é claro, dispositivos físicos e equipamentos”³ (DOSI, 1982,

3 Traduzido do original em inglês.

p.151; DOSI, 1984, p.14). Ou seja, o conhecimento tecnológico pode estar, ou não⁴, incorporado em um dispositivo físico existente, ou ser representado, em determinada dimensão, por este⁵. Adicionalmente, argumenta que a visão de tecnologia a ser explorada difere do conceito de tecnologia como sendo informação amplamente aplicável e de fácil reprodução e reutilização, (DOSI, 1988), caracterizando o processo inovativo como sendo o resultado de uma imersão gratuita em um estoque de conhecimento tecnológico de livre acesso. Antes, o autor defende que a informação é um subconjunto da tecnologia, ou do conhecimento tecnológico, que engloba outro elemento dependente de habilidades e competências que crescem cumulativamente. “Em cada tecnologia existem elementos de conhecimento tácito e específico que não são e não podem ser escritos na forma de um manual de instrução, e não podem, dessa forma, ser inteiramente difundido, nem em forma de informação pública nem privada” (DOSI, 1988, p.1131)⁶. A evolução da tecnologia, por sua vez, se daria então pela recombinação de conhecimentos, que seriam os *building blocks* da tecnologia, sendo possibilitada pelo acúmulo crescente destes através do processo inovativo (ARTHUR, 2009). Ou seja, a tecnologia evolui através da inovação tecnológica.

Dosi (1988 e 1988a) argumenta que a inovação tecnológica, em geral, envolve a solução de problemas, satisfazendo, ao mesmo tempo, alguns requerimentos de custo e comercialidade, implicando em um conjunto de conhecimento de diversos tipos. Esses problemas tecnológicos não podem ser solucionados a partir da informação disponível, ou seja, requerem soluções inovativas, demandando descoberta e criação. Estas soluções inovativas se apoiam sobre uma base de conhecimento composta por conhecimento formal (como as ciências naturais), informações advindas de experiências prévias e capacidades específicas e não codificadas (conhecimento específico e tácito). Ou seja, o conjunto de conhecimentos subjacente à solução dos problemas tecnológicos, em geral, apresenta os seguintes aspectos: universalidade (conhecimento científico ou conhecimento relacionado à princípios aplicativos abrangentes e bem conhecidos) ou especificidade (conhecimento específico a um particular ‘jeito de fazer as coisas’, a experiência do produtor e/ou do usuário); codificado ou tácito (conhecimento disponível em manuais, livros e artigos versus conhecimento adquirido na prática); e público

4 “[...] uma parte 'desencarnada' da tecnologia consiste em conhecimento específico, experiência de tentativas passadas e soluções tecnológicas passadas, juntamente com o conhecimento e as realizações do estado-da-arte” (DOSI, 1984, p.14). [Traduzido originalmente do inglês]

5 Dosi (1984, p.14): “Dispositivos físicos existentes incorporam as conquistas no desenvolvimento de uma tecnologia em uma atividade solucionadora de problemas definida”. [Traduzido originalmente do inglês]

6 Traduzido originalmente do inglês.

versus privado (publicações versus conhecimento tácito ou protegido por alguma medida de apropriabilidade), (DOSI, 1988a). Assim, a inovação tecnológica irá se utilizar de algum subconjunto do conhecimento publicamente disponível, porém, implicando, obrigatoriamente, o uso e desenvolvimento concomitante de conhecimento específico e, em geral, parcialmente privado. Dosi (1988) salienta que uma importante caracterização das diferentes tecnologias deriva do grau de publicidade e universalidade versus tacitividade e especificidade (geralmente um conhecimento gerado dentro das unidades inovativas) de suas bases de conhecimento. Pontua ainda que qualquer que seja a base de conhecimento sobre a qual a inovação se realiza, cada atividade de solução de problemas implica o desenvolvimento e refinamento de “modelos” e procedimentos específicos (p.1127), ou seja, a inovação se caracteriza por seu um processo inerentemente acumulativo.

Dosi (1988 e 1998a) ainda define que um outro aspecto da característica pública da tecnologia, para além dos mencionados acima, advém da existência do que chamou de “interdependências não negociadas”, existentes entre firmas, setores e tecnologias, que, em geral, assumem a forma de complementariedades tecnológicas, sinergias e fluxos de estímulos e restrições, os quais não correspondem totalmente ao fluxo de mercadorias, ou o extrapolam. Ou seja, não envolvem necessariamente alguma transição econômica. Estes aspectos representariam um conjunto de externalidades tecnológicas.

Em outras palavras, os gargalos e as oportunidades tecnológicas (Rosenberg, 1976), as experiências e qualificações incorporadas em pessoas e organizações, as capacitações e “memórias”, transbordando de uma atividade econômica a outra, tendem a organizar *condições contextuais* que (a) são específicas aos países, às regiões ou mesmo às firmas; (b) são um ingrediente fundamental no processo inovativo; e (c) como tais, determinam diferentes incentivos/estímulos/restrições à inovação, para qualquer dada série de sinais estritamente econômicos. (DOSI, 1998a, p.226)⁷

Enxergando certa similaridade nas definições e procedimentos entre a ciência e a tecnologia, Dosi (1982) remete à elaboração de Thomas Kuhn que definiu o “paradigma científico” (ou programas de pesquisa científica) como sendo o conjunto de conhecimento disponível, mais avançado e com maiores perspectivas de avanço futuro, que define um conjunto de programas e linhas de ação para novas pesquisas (KUHN, 2017: 1962). Ou seja, “uma perspectiva que define os problemas relevantes, um modelo e um padrão de investigação”⁸ (DOSI, 1982, p.152).

7 Traduzido do original em inglês.

8 Traduzido do original em inglês.

Assim, Dosi (1982, 1984, 1988, 1988a), por analogia, define paradigma tecnológico como um “panorama, um conjunto de procedimentos, uma definição dos problemas relevantes e do conhecimento específico relacionado as suas soluções” (DOSI, 1982, p.148). Ou seja, determinados problemas tecnológicos encontrariam soluções dentro de um conjunto delimitado de conhecimentos relevantes que compõe o paradigma tecnológico. Ou ainda, o paradigma “como um modelo e um padrão de solução de problemas tecnológicos selecionados, baseado em princípios selecionados derivados das ciências naturais e em tecnologias materiais selecionadas”⁹ (DOSI, 1982, p.152). Assim, o paradigma tecnológico define a área de atuação – ou de pesquisa -, os problemas, os procedimentos e os desafios – ou objetivos – a resolver. Dosi (1988) traz ainda a seguinte definição de paradigma tecnológico: “um ‘padrão’ de solução de problemas tecno-econômicos selecionados, baseado em princípios altamente selecionados derivados das ciências naturais, juntamente com regras específicas objetivando adquirir novo conhecimento e salvaguardá-lo, quando possível, da rápida difusão para os competidores¹⁰” (p.1127).

Admitir que a inovação é uma atividade de solução de problemas, delimitada pelos paradigmas tecnológicos, conseqüentemente estabelece a natureza acumulativa do processo inovativo, ou seja, a acumulação crescente de conhecimento tecnológico que torna as tecnologias e o processo inovativo cada vez mais complexos. “Uma implicação crucial da forma geral paradigmática do conhecimento tecnológico é que atividades inovativas são fortemente seletivas, focalizada em direções precisas, e cumulativa na aquisição de capacidades para solução de problemas¹¹”, (DOSI, 1988, p.1128). Alguns exemplos de paradigmas tecnológicos bem conhecidos são o motor a combustão interna, petroquímicos e semicondutores. Ainda, Dosi (1988) argumenta que um paradigma tecnológico é tanto um exemplar (um artefato a ser desenvolvido e melhorado, cada um com suas características tecno-econômicas particulares) quanto um conjunto de heurísticas (para onde ir daqui? Onde procurar? Em qual tipo de conhecimento se basear?). Em outro trabalho, Dosi ainda afirma: “paradigmas tecnológicos definem as oportunidades tecnológicas para mais inovações e alguns procedimentos básicos sobre como explorá-las¹²” (DOSI, 1988a, p.225). Em Dosi (1984) nota-se ainda uma referência aos paradigmas tecnológicos como sendo um cluster de tecnologias.

9 Traduzido do texto original em inglês.

10 Traduzido do texto original em inglês.

11 Traduzido do texto original em inglês.

12 Traduzido do texto original em inglês.

Dosi (1988 e 1988a) salienta a existência de paradigmas com diferentes níveis de generalidade quando se observa os padrões de mudança técnica. Ademais, cada tecnologia apresentaria procedimentos, competências e heurísticas específicas em seu processo de busca, ou seja, cada paradigma tecnológico envolveria uma “tecnologia da mudança técnica” específica (p.1128). À medida que o paradigma é explorado economicamente e reproduzido, novas instituições são desenvolvidas para assegurar sua manutenção. Modelos prototípicos de solução de problemas se tornam a visão geral, servindo como sinalizadores tecnológico ou “*focusing device*”. Outro ponto importante trata da forte incerteza característica dos estágios pré-paradigmáticos, como denomina Dosi (1988), da mudança técnica, períodos de elevada atividade exploratória – períodos de “exploração extraordinária - onde há incertezas sobre os resultados práticos da busca inovativa e, também, quantos aos princípios científicos e tecnológicos, bem como quanto aos procedimentos “solucionadores de problemas” sobre os quais os avanços tecnológicos podem ser baseados. Assim, “quando um paradigma tecnológico é estabelecido, traz consigo uma redução da incerteza, no sentido de que ele focaliza as direções da busca e cria as bases para formação de expectativas tecnológicas e de mercado com mais segurança¹³” (DOSI, 1988, p.1134). “[...] um paradigma tecnológico incorpora fortes prescrições sobre as direções das mudanças técnicas a serem buscadas e aquelas a serem negligenciadas¹⁴” (DOSI, 1984, p.15).

Uma vez estabelecido, cada paradigma tecnológico define seu próprio conceito de progresso tecnológico, baseado em seus trade-offs tecnológicos e econômicos específicos. Assim, Dosi (1982) define uma trajetória tecnológica como sendo “o padrão de atividade normal solucionadora de problema (ou de progresso) no terreno do paradigma tecnológico¹⁵” (p.152). Ou seja, é a direção do avanço tecnológico dentro de um determinado paradigma. Isso acontece porque cada paradigma tecnológico “incorpora fortes prescrições sobre as direções de mudança técnica a perseguir e aquelas a negligenciar¹⁶” (p.152). Dessa forma, existe um poderoso efeito exclusão por um lado e uma definição de progresso do outro. Segundo Dosi (1982 e 1984) a identificação de um paradigma tecnológico estaria relacionada com as questões gerais sobre as quais se debruça, às tecnologias materiais que seleciona, às propriedades físico-químicas que explora, às dimensões econômicas e tecnológicas e trade-offs que foca (p.153). E, uma vez

13 Traduzido do texto original em inglês.

14 Traduzido do texto original em inglês.

15 Traduzido do texto original em inglês.

16 Traduzido do texto original em inglês.

definidas estas dimensões econômicas e tecnológicas, é possível obter a ideia de progresso como uma melhora dos trade-offs relacionados à essas dimensões. Trajetória tecnológica seria, assim, a atividade de progresso tecnológico ao longo dos trade-offs econômicos e tecnológicos definido por um paradigma tecnológico (DOSI, 1988, p.1128). A existência dessas “avenidas de inovação” específicas são uma característica generalizada dos padrões observados de mudança técnica. Evidências empíricas sugerem que o maior impulso para inovação deriva dos desequilíbrios entre as dimensões técnicas que caracterizam uma “trajetória” (DOSI, 1988, p. 1129). As trajetórias encarnariam, assim, uma descrição *ex-post* dos padrões da mudança técnica, como também a base da heurística sobre para onde ir a partir do atual ponto de evolução tecnológica (DOSI, 1988, p.1134).

Dosi (1982 e 1984) define ainda algumas características das trajetórias tecnológicas que são estabelecidas de acordo com o paradigma tecnológico vigente: variação no grau de generalidade e poder das trajetórias – poder, aqui, sendo definido como a capacidade de excluir o maior número possível de outras tecnologias -; existência de complementariedades entre trajetórias; definição da fronteira tecnológica, com base nas dimensões tecnológicas e econômicas, como sendo o maior nível alcançado dentro de uma trajetória tecnológica; progresso ao longo de uma trajetória tecnológica possui efeitos acumulativos e a possibilidade de futuros avanços dependem da posição na trajetória e a fronteira tecnológica; dificuldade na mudança de uma trajetória para outra (especialmente se a trajetória for poderosa), ou seja, sempre que se altera o paradigma tecnológico se faz necessário iniciar numa nova trajetória do início; impossibilidade de se acessar e comparar a priori a superioridade de uma trajetória sobre outra, o que confere a natureza de incerteza ao processo inovativo, dentre outros fatores (DOSI, 1982, p.154). Relacionado a isto, o progresso tecnológico ao longo de qualquer trajetória está ligado (a) ao desenvolvimento de infraestrutura específica; (b) a economias de escala sistêmicas; (c) a tecnologias complementares; e (d) normas técnicas particulares que se alimentam positivamente dos padrões específicos de inovação. (DOSI, 1988, p.1146)

Sobre como um determinado paradigma tecnológico surge, Dosi (1988) argumenta que a fonte de novos paradigmas, cada vez mais, depende de avanços fundamentais na ciência e nas tecnologias “gerais”. E o processo pelo qual um paradigma é selecionado em detrimento de outros possíveis se baseia nas forças econômicas juntamente com fatores institucionais e sociais operando como um dispositivo de seleção. Primeiro, considerando o processo de gênese de novos paradigmas: a emergência de novos potenciais paradigmas que concerne na geração de

oportunidades nocionais de inovações radicais envolvendo novas bases de conhecimento, novas heurísticas de pesquisa, novos designs dominantes, (DOSI, ORSENIGO, LABINI, 2005). Ou seja, dentre as inúmeras possibilidades disponíveis através dos desenvolvimentos científicos¹⁷ e tecnológicos, e após uma certa seleção própria interna à essas esferas¹⁸, os fatores econômicos, institucionais e sociais desempenham um papel preponderante na definição dos caminhos a seguir. Uma vez que esse caminho tenha sido selecionado, ou seja, que o paradigma tenha se estabelecido, ele define as direções nas quais as atividades se concentrarão¹⁹, ou seja, suas “trajetórias naturais de progresso técnico”²⁰, que são as trajetórias tecnológicas definidas anteriormente nesta seção. Uma trajetória tecnológica seria assim um cluster de possíveis direções tecnológicas (ou oportunidades tecnológicas) cujos limites são definidos pela natureza do paradigma atuante.

A princípio, a seleção desempenhada pelos fatores econômicos, institucionais e sociais se circunscrevem à aspectos gerais como viabilidade, comercialização e lucratividade. No entanto, essa fase inicial de seleção resulta na existência de muitos paradigmas tecnológicos possíveis. Assim, ocorre o refinamento do mecanismo de seleção através de variáveis mais específicas como *focussing device* de direções definidas do desenvolvimento tecnológico, a saber: interesse econômico das organizações envolvidas no P&D dessas novas áreas tecnológicas, sua história tecnológica, os campos de expertise, variáveis institucionais como agências públicas, forças armadas etc. Além disso, o papel desempenhado por forças políticas ou públicas na definição de uma trajetória tecnológica específica, a exemplo das corridas espaciais. Outro critério de seleção: capacidade de economia de custos da nova tecnologia, em especial, de custos do fator

17 Dosi (1988) salienta que o conhecimento científico, a partir do século XX, desempenhou um papel crucial na abertura de novas possibilidades de grandes avanços tecnológicos e que a emergência novos paradigmas tecnológicos expressivos tem frequentemente dependido diretamente e diretamente ligado com grandes descobertas científicas, como é o caso da química sintética, por exemplo. “[...] a geração e utilização de parte do conhecimento científico é interna e, geralmente, um condição necessária ao desenvolvimento de novos paradigmas tecnológicos” (DOSI, 1988, p.1136). O progresso do conhecimento científico geral, então, criaria uma ampla gama de paradigmas tecnológicos potenciais.

18 Dosi (1984 e 1988) estabelece que o processo de seleção dentre as oportunidades criadas pela ciência depende em geral da natureza e interesses dos instituições que fazem a ponte entre as pesquisa pura e as aplicações econômicas, fatores estritamente institucionais, dos mecanismos de tentativa e erro da exploração de novas tecnologias e dos critérios de seleção dos mercados e ou dos requerimentos tecno-econômicos dos primeiros usuários.

19 Vale ainda ressaltar que a partir do estabelecimento de um paradigma tecnológico, os objetivos e heurísticas da busca tecnológica tende a divergir daqueles da pesquisa científica. No entanto, atividades de pesquisa aplicada se tornam parte da busca tecnológica ao longo das trajetórias definidas pelo paradigma em questão, Dosi (1988).

20 Conceito apresentado por Nelson e Winter (1977).

trabalho²¹. Os padrões de conflito industrial e social também operam como critério de seleção de novos paradigmas. Existência de uma relação entre padrões de desenvolvimento social e paradigmas tecnológicos selecionados.

Assim, o ambiente econômico e social afeta o desenvolvimento tecnológico de duas maneiras, primeiro selecionando a 'direção da mutação' (ou seja, selecionando o paradigma tecnológico) e, em seguida, selecionando entre as mutações, de uma maneira mais darwiniana (ou seja, a seleção ex post entre as tentativas e erros 'schumpeterianas') (DOSI, 1982, p. 156).

[...] fazer uma distinção lógica entre o processo de busca e seleção de novos paradigmas tecnológicos e o progresso técnico ao longo de um caminho definido. Novas tecnologias são selecionadas por meio de uma interação complexa entre alguns fatores econômicos fundamentais (busca de novas oportunidades de lucro e de novos mercados, tendência para fatores institucionais de custo poderosos (os interesses e a estrutura das empresas existentes, os efeitos das agências governamentais, etc.). A mudança técnica ao longo de caminhos tecnológicos estabelecidos, ao contrário, torna-se mais endógena ao mecanismo econômico 'normal' (DOSI, 1982, p.157)²².

Os novos paradigmas remodelam os padrões de oportunidades de progresso técnico, além de, em geral, difundir seus efeitos para além do setor de origem, promovendo novas fontes de oportunidades via fluxos de insumo-produto e complementariedades tecnológicas. Ademais, novos paradigmas tecnológicos geralmente impedem, direta ou indiretamente, via seus efeitos sobre os paradigmas precedentes, o estabelecimento de retornos decrescentes no processo de busca por inovações. Vale ainda ressaltar, como fez Dosi (1988), que o aparecimento de novos paradigmas tecnológicos é desigualmente distribuído entre setores, assim como o são as oportunidades e as competências tecnológicas para inovar. Estas últimas dependem da natureza das atividades setoriais, da distância tecnológica do núcleo revolucionário onde novos paradigmas são originados e da base de conhecimento que sustenta a inovação em cada setor.

As definições de paradigma tecnológicos e trajetórias tecnológicas ajudam a estabelecer uma distinção entre a mudança técnica versus progresso técnico, entre o papel da continuidade e da descontinuidade na mudança técnica, entre inovações incrementais versus inovações radicais, ao definir o que é progresso tecnológico normal e a emergência de novos paradigmas. Dosi (1988) faz uma distinção entre o que chamou de “progresso técnico normal” – concernente aos processos de inovação internos aos limites de um dado paradigma tecnológico – e “progresso técnico extraordinário” – que está associado ao surgimento de novos paradigmas. O conjunto

21 Rosenberg (1976) argumenta que: “*the ultimate incentives are economic in nature; but economic incentives to reduce cost always exist in business operations, and precisely because such incentives are so diffuse and general they do not explain very much in terms of the particular sequence and timing of innovative activity*” (p.110).

22 Traduzido do texto original em inglês.

de possíveis trajetórias, relacionado ao primeiro termo, é limitado e circunscrito pelas regras e pelos imperativos tecnológicos de cada tecnologia, que, no curto prazo, parecem ser invariantes às condições de mercado. Estas, parecem estar mais relacionadas, num horizonte temporal mais amplo, com a condução da busca tecnológica na direção de novos paradigmas: “avanços científicos, juntamente com vários fatores institucionais, contribuem para determinar o momento e a natureza de novos paradigmas, enquanto os mercados atuam como importantes ambientes de seleção”²³ (DOSI, 1984).

[... cada paradigma] restringe tanto as oportunidades de progresso técnico quanto os limites dentro dos quais os “efeitos indutores” poderão ser exercidos pelo mercado, enquanto as condições de apropriabilidade motivam os agentes econômicos a explorarem essas oportunidades tecnológicas como um dispositivo competitivo gerador de renda. Finalmente, a evolução do ambiente econômico, em um prazo mais longo, é instrumental na seleção de novos paradigmas tecnológicos e, assim, na seleção a longo prazo das direções fundamentais e dos procedimentos de busca inovativa (DOSI, 1988, p. 1142)²⁴.

Por fim, quando o paradigma vigente perde a capacidade de solucionar os problemas tecnológicos de forma satisfatória ou a capacidade de ditar o modo de operação da solução desses problemas, ele é sucedido por outro que apresente melhores perspectivas. O surgimento de novos paradigmas potenciais acontece ao longo da acumulação de conhecimento proporcionada pelos paradigmas vigentes. Vale ainda ressaltar que, em geral, nota-se a sobreposições de paradigmas tecnológicos e que mudanças nestes, recorrentemente, estão associadas às crises e aos ciclos do sistema capitalista.

Brian Arthur (2009), escrevendo sobre o processo de evolução da tecnologia, faz uma contribuição para o entendimento dos conceitos de paradigmas e trajetórias tecnológicas, ao defender a ideia de que a tecnologia não surge do nada, mas, antes, é o resultado de um processo evolutivo onde novas tecnologias surgem a partir de outras tecnologias já existentes. Um processo de seleção que envolve fatores científicos, tecnológicos, econômicos, institucionais, sociais e complexos mecanismo de feedbacks (ARTHUR, 2009). O processo de evolução da tecnologia, via inovação, se daria a partir de novas combinações de tecnologias existentes, como *building blocks*, e do uso do que chamou de fenômenos naturais para atender objetivos específicos. Assim, também enfatiza o caráter cumulativo da tecnologia, onde tecnologia gera mais tecnologia, e cria a ideia de que se pode traçar uma linha genética, ou linha de

23 Traduzido do original em inglês.

24 Traduzido do original em inglês.

descendência, das tecnologias existentes, já que existe um mecanismo de hereditariedade na evolução. Ademais, afirma o potencial de surgimento de cada vez mais tecnologia, dado que o conjunto de tecnologias a serem combinadas só aumenta. Porém, as combinações não são feitas randomicamente, existe uma lógica governando a dinâmica inovativa, pois a tecnologia seria uma combinação de componentes, ou partes, para algum propósito, sendo organizada ao redor de um conceito ou princípio geral, que, por sua vez, constitui o método ou a ideia essencial que permite que ela funcione. Arthur (2009) também ressalta que a tecnologia pode ser pensada tanto no sentido mais abstrato quanto no específico.

Para ser trazido à realidade física, um princípio precisa ser expresso na forma de componentes físicos. Na prática, isso significa que uma tecnologia consiste em um conjunto principal: uma espinha dorsal geral do dispositivo ou método que executa seu princípio básico. Esta espinha dorsal é apoiada por outros conjuntos para cuidar de seu funcionamento, regular sua função, alimentá-lo de energia e realizar outras tarefas subsidiárias. Portanto, a estrutura primária de uma tecnologia consiste em um conjunto principal que realiza sua função básica mais um conjunto de subconjuntos que suportam este (ARTHUR, 2009, p.28)²⁵.

A tecnologia, pensada individualmente ou em grupo, seria composta por uma combinação hierárquica de *building blocks*, constituindo seus subsistemas que interagem de forma a melhor servir ao propósito geral. Dessa forma, cada parte precisa se ajustar as demais, sendo combinadas para isso, de forma a alcançar o princípio a que se destina o sistema como um todo, resultando em um processo que lida com vários trade-off. A partir dessas definições, pode-se traçar uma analogia com os conceitos de paradigmas e trajetórias tecnológicas. Assim como na definição de Dosi (1982, 1984, 1988), Arthur (2009) também trabalha com uma definição de inovação como sendo uma atividade de solução de problemas e como o processo pelo qual a tecnologia avança. Dada a existência de um propósito – ou um problema –, a atividade inovativa se traduz na junção de um fenômeno (ou efeito, geralmente concebido pela atividade científica) com um propósito, para solucioná-lo, nesse interim, surge o princípio base que governará o funcionamento da nova tecnologia. Tal princípio, ou tecnologia, se traduz na solução do problema. Essa inovação seguiria um ciclo: surgimento do novo princípio, desenvolvimento (através do alcance de obstáculos e sua superação) e complexificação, maturidade, persistência através de um processo denominado “esticamento” para tentar atender às novas demandas e circunstâncias, alcance dos últimos limites e sofisticções e, por fim, ceder e ser substituído pelo novo princípio ou tecnologia.

25 Traduzido do original em inglês.

Quando da definição desse ciclo, Arthur (2009) enxerga uma clara analogia com os ciclos de desenvolvimento das teorias científicas, estabelecidas por Thomas Kuhn. Isso se deve ao fato de que as teorias científicas, assim como a tecnologia, são definidas como sendo “*purposed systems*”:

O ciclo de Kuhn começa quando um novo modelo teórico trabalhando em um novo princípio (ele chama isso de paradigma) substitui um antigo. O novo paradigma é então elaborado, aplicado a muitos exemplares, torna-se aceito e é posteriormente elaborado em um processo que Kuhn chama de ciência normal. Com o tempo, exemplos que não se encaixam no paradigma básico - anomalias - aumentam. O paradigma é ampliado para acomodá-los, mas torna-se cada vez mais tenso à medida que novas anomalias se acumulam. Ele entra em colapso quando e somente quando um novo e mais satisfatório conjunto de explicações - um novo paradigma - chega (ARTHUR, 2009, p.112)²⁶.

Assim, o princípio base da tecnologia definiria o paradigma tecnológico. Este, que pode ser entendido como compondo um novo domínio tecnológico, em geral emerge ao redor de uma tecnologia central, ou uma tecnologia capacitadora, que tem a característica de ser capaz de atrair o interesse de outros agentes, criar companhias, atrair investimentos e o interesse dos jornalistas, ou seja, o novo campo começa a se tornar preponderante. Arthur (2009) argumenta que um evento comum nesse estágio é a valorização rápida das ações das novas firmas e que muitas delas adentram o mercado sem muito mais a oferecer do que uma ideia, somente apresentando um produto comercializável após algum tempo passado da sua primeira oferta pública de ações. Esse foi o caso, por exemplo, da IPO da Netscape, que viu suas ações valorizarem num tempo recorde e sem um produto definido a oferecer. No entanto, a tecnologia da *World Wide Web* já atraía considerável atenção a ponto de permitir tal valorização.

Outra característica importante dos domínios tecnológicos é sua capacidade de mutação, aplicando o mesmo princípio básico em outros campos, conferindo uma qualidade de organismo vivo à um corpo de tecnologias. Assim, haveria sempre a criação de sub colônias de tecnologias que se desenvolveriam de forma independente e com diferentes características. O que pode ser visto no atual desenvolvimento de campos tecnológicos como os da inteligência artificial, plataformas, blockchains e Internet das coisas, remetendo à essa característica de mutação do paradigma inaugurado pela WWW. Além disso, Arthur (2009) argumenta que não são apenas os grandes domínios tecnológicos que geram revoluções, mas estas também ocorrem

26 Tradução do original em inglês.

de forma sobreposta, interativa e simultânea, alterando a economia, na medida em que domínios de menor escala surgem. O que gera uma das hipóteses de análise da WWW que, mesmo não configurando uma revolução tecnológica nos termos definidos por Carlota Perez (2002 e 2004), dificilmente não se enquadraria numa definição de revolução dada sua capacidade de transformação da economia, sociedade e instituições.

Uma revolução não chega completamente até que organizemos nossas atividades - nossos negócios e procedimentos comerciais - em torno de suas tecnologias e até que essas tecnologias se adaptem a nós. Para que isso aconteça, o novo domínio deve reunir adeptos e prestígio. Deve encontrar propósitos e usos. Suas tecnologias centrais devem resolver certos obstáculos e preencher certas lacunas em seu conjunto de componentes. Deve desenvolver tecnologias que o suportem e conectem com as tecnologias que o utilizam. Ele deve compreender seus fenômenos básicos e desenvolver a teoria por trás deles. Os mercados devem ser encontrados e as estruturas existentes da economia devem ser redesenhadas para fazer uso do novo domínio. E a velha dispensação deve reconhecer o novo domínio e se familiarizar com suas práticas inerentes, o que significa que os engenheiros que comandam a gramática do antigo precisam se reequipar para o novo. Eles não fazem isso levemente. Tudo isso deve ser mediado pelas finanças, pelas instituições, pela gestão, pelas políticas governamentais e pela disponibilidade de pessoas qualificadas no novo domínio (ARTHUR, 2009, p.125)²⁷.

Vale a pena uma nota sobre a diferença entre os conceitos de paradigma tecnológico proposto por Dosi (1982, 1984, 1988, e 1988a), amplamente descrito anteriormente, e o de paradigma tecno-econômico, proposto inicialmente por Freeman e Perez (1986 e 1988)²⁸ e posteriormente sofisticado por Perez (2002). O conceito de paradigma tecno-econômico refere-se aos clusters mais amplos de paradigmas tecnológicos, às meta-rotinas, aos clusters de cluster de inovações capazes de alterar a atividade econômica em sua totalidade, sendo iniciados por revoluções tecnológicas, e correspondendo aos grandes surtos ou ondas (no sentido schumpeteriano) de desenvolvimento econômico. Segundo essa abordagem, são identificadas cinco revoluções tecnológicas, a partir da primeira revolução industrial, com duração aproximada de 50 anos, com o estabelecimento do respectivo paradigma tecno-econômico a guiar o desenvolvimento tecnológico e econômico subsequente. Esse conceito se distingue do conceito proposto por Dosi, principalmente no que tange ao alcance ou escopo da transformação abarcada.

27 Tradução do original em inglês.

28 Diferenciação dos tipos de mudança técnica proposta por Freeman e Perez (1988) através da taxonomia sugerida: inovações radicais, inovações incrementais, sistemas tecnológicos e paradigmas tecno-econômicos.

Por fim, é interessante tecer uma nota sobre a atualidade e força analítica desse arcabouço teórico desenvolvido a partir do trabalho de Dosi (1982). Em uma edição especial, o periódico *Industrial e Corporate Change* de 2008 se dedica ao estudo do impacto da evolução dessas ideias, ou como afirmam: “colocamos essas ideias em seu contexto histórico e avaliamos seu impacto passado, presente e futuro na compreensão acadêmica, prática gerencial e política governamental no que se refere à ciência, tecnologia e inovação²⁹” (VON TUNZELMANN, MALERBA, NIGHTINGALE E METCALFE, 2008, p.467). Os autores realizaram um estudo bibliométrico para avaliar o impacto acadêmico do artigo original, Dosi (1982), concluindo que este é um dos artigos mais citados dentro da literatura da economia da mudança técnica, além de transbordar para outros campos além da economia. Salientam, ainda, uma consequente introdução de certo rigor científico nas discussões sobre as tendências na tecnologia a partir do trabalho de Dosi (1982). No entanto, o argumento dos autores se estende para além dessa observada importância, defendendo que as ideias desenvolvidas por Dosi (1982) ainda possuem considerável potencial analítico não explorado.

1.2. General Purpose Technologies.

O primeiro uso do termo *General Purpose Technology* (GPT) é atribuído aos textos, pioneiros da literatura de GPTs, de Bresnahan e Trajtenberg (1992 e 1995). A partir do reconhecimento de que, de uma perspectiva histórica, algumas poucas invenções parecem ter desempenhado um papel chave para a dinâmica capitalista³⁰, busca-se definir a natureza de tais tecnologias possuidoras do potencial para afetar o crescimento econômico ao longo de um prolongado período de tempo. Bresnahan (2010) argumenta que, de fato, a motivação original para a ideia de GPT, na análise econômica da mudança técnica, surgiu, em parte, da história do crescimento econômico quando do estudo de algumas tecnologias chaves – como o motor a vapor e a eletricidade – e da determinação de seu papel central no processo de crescimento. Ainda de acordo com Bresnahan (2010), para além dessa motivação, localizada na macroeconomia do crescimento, outras vieram em seguida, como a análise da microeconomia da mudança técnica com foco na diferenciação entre inovações de diferentes tipos: “os incentivos e informações relacionados à invenção das próprias GPTs, por exemplo, podem diferir daqueles relacionados à invenção das aplicações” (BRESNAHAN, 2010, p.763). Ademais, outra motivação seria uma

29 Tradução do original em inglês.

30 Os autores citam o motor a vapor, o sistema de fábricas, eletricidade e os semicondutores.

intersecção entre essas questões macro e micro, na tentativa de entender os links entre crescimento econômico, estruturas de incentivos e informacionais relacionadas à uma invenção particular e sua aplicação em setores e usos específicos.

Cada um desses objetivos recebeu atenção considerável, com uma combinação diferente de métodos teóricos, empíricos e históricos. Houve um grande progresso, embora questões permaneçam em aberto sobre como testar as idéias mais interessantes e importantes. Algumas críticas à abordagem também surgiram. Isso torna este um bom estágio para revisar os objetivos originais conforme alterados por novas (BRESNAHAN, 2010, p.763)³¹.

Estas tecnologias, classificadas como GPTs, possuiriam três características conceituais chaves: generalidade, dinamismo tecnológico e complementariedades inovadoras (BRESNAHAN E TRAJTENBERG, 1995; LIPSEY, BEKAR e CARLAW, 1998; HELPMAN, 1998; BRESNAHAN, 2010). Tais características seriam as responsáveis por proporcionar aumentos de produtividade generalizados ao logo da economia, determinando, em última instância, o crescimento econômico. Ou seja, à medida que a GPT evolui e avança, se espalhando por toda a economia, promove a aumenta ganhos de produtividade generalizados (BRESNAHAN E TRAJTENBERG, 1995).

A generalidade trata da capacidade da tecnologia de dispor de aplicabilidade nos mais diversos e numerosos setores da economia (denominados setores de aplicação), ou apresentar disponibilidade para uso generalizado, ou ainda, capacidade de oferecer uma função genérica, (BRESNAHAN E TRAJTENBERG, 1995; ROSENBERG, 1998; LIPSEY, BEKAR e CARLAW, 1998; HELPMAN, 1998; BRESNAHAN, 2010). O motor elétrico, por exemplo, demonstrou aplicabilidade em máquinas industriais, carros, eletrodomésticos, entre outros. Afirma-se ainda que GPTs com característica exacerbada de generalidade permitem inovações muito diversas nos setores de aplicação. O dinamismo tecnológico, por sua vez, se refere ao potencial que a GPT demonstra para suportar esforços inovadores e de aprendizagem contínuos, o que possibilita grandes elevações de eficiência da GPT ao longo do tempo, ou seja, sua capacidade de aprimoramento técnico contínuo, (BRESNAHAN E TRAJTENBERG, 1995; LIPSEY, BEKAR E CARLAW, 1998; BRESNAHAN, 2010).

31 Traduzido do original em inglês.

As complementariedades inovadoras se relacionam ao fenômeno da GPT desempenhar o papel de “*enabling technologies*”, possibilitando novas oportunidades, ao invés de oferecer soluções finais completas, ou seja, uma tecnologia que possibilita novas fontes de invenções e inovações nos setores de aplicação (BRESNAHAN E TRAJTENBERG, 1995; LIPSEY, BEKAR E CARLAW, 1998; BRESNAHAN, 2010). Dessa forma, a produtividade da atividade de pesquisa e desenvolvimento (P&D) em um setor de aplicação aumenta como consequência da inovação na tecnologia GPT. Assim, elevam-se os retornos às inovações em cada setor de aplicação e vice-versa, reduz-se o custo nos setores de aplicação e percebe-se melhoras na qualidade e variedade dos produtos dos setores etc., (BRESNAHAN, 2010). Essas complementaridades ampliam os efeitos da inovação na GPT e ajudam a propagá-los por toda a economia. Fazer com que o "conceito geral" funcione em qualquer situação específica requer maiores inovações complementares “e, geralmente, uma grande quantidade de ingenuidade” (BRESNAHAN E TRAJTENBERG, 1995). Paralelamente ao surgimento de novas aplicações, a GPT promove mudanças contínuas nos setores já existentes. De maneira mais geral, adaptar e adotar a GPT para diferentes setores é uma atividade inovadora.

Essas inovações complementares, segundo Bresnahan e Trajtenberg (1995), criam oportunidades e desafios para o crescimento via avanço técnico. As oportunidades derivam de desenvolvimentos de aplicações ou usos da tecnologia GPT numa ampla variedade de setores, aumentando o retorno para novos avanços no próprio setor da GPT, que, por sua vez, são capazes de sofisticar a tecnologia inicial e levar a novas possibilidades de aplicações. Essa dinâmica de feedbacks positivos reforçariam uma rápida mudança técnica e o consequente crescimento econômico, via geração de ganhos de produtividade em todos os setores, e aprofundaria os efeitos da tecnologia inicial.

Já as dificuldades, advindas da dinâmica das inovações complementares, se traduziriam, segundo Bresnahan e Trajtenberg (1995), na dispersão dessas atividades ao longo de toda a economia, o que tornaria difícil a coordenação e a promoção dos incentivos inovativos adequados para os setores da GPT e os de aplicação. Essas dificuldades de coordenação resultariam das características essenciais do processo de criação de novos conhecimentos: incerteza, informação assimétrica, intervalos de tempo e sequencialidade. Portanto, a coordenação nesse contexto exigiria o alinhamento dos incentivos dos agentes localizados longe um do outro ao longo do tempo e nas dimensões da tecnologia. Como as GPTs estão

conectadas, por definição, a amplos segmentos da economia, falhas de coordenação dessa natureza podem ter consequências de longo alcance para o crescimento econômico.

Em suma, a formalização inicial de Bresnahan e Trajtenberg (1995) considerava duas hipóteses técnicas chaves, a generalidade de propósito e as complementariedades inovadoras, que se traduziam em duas externalidades distintas: externalidade vertical entre a GPT e cada setor de aplicação e a externalidade horizontal entre os setores de aplicação.

No entanto, é importante salientar que para além dessas características, acima citadas, que compõe a definição da estrutura básica da GPT, outros elementos são de crucial relevância para o entendimento da dinâmica de cada GPT em particular, podendo variar entre estas tecnologias, como, por exemplo, a forma como é ofertada a inovação da GPT e dos setores de aplicação, suas características econômicas e a natureza da interação entre os setores inovadores da GPT e dos setores de aplicação (BRESNAHAN, 2010). A GPT pode estar, ou não, incorporada a um bem e serviço que é adquirido pelos setores de aplicação, protegida por patentes e segredos comerciais, ou ser de domínio público. As estruturas organizacionais e a forma de relação estabelecida entre a GPT e a invenção nos setores de aplicação também podem apresentar formas variadas e distintas: a GPT possibilitando oportunidades técnicas, ou de mercados, para os setores de aplicação, ou apenas transferindo um bem e serviço a ser utilizado pelos setores de aplicação. Bresnahan (2010) salienta que “as implicações do modelo específico de interação não são as mesmas da estrutura básica da GPT” (p. 766)³². As abstrações do modelo teórico de definição da GPT podem esconder diferenças de evolução e efeitos macroeconômicos de cada tecnologia em específico (BEKAR, CARLAW e LIPSEY, 2018) e como esses macrofenômenos estão profundamente enraizados nos micro fundamentos da mudança técnica.

Um exemplo de uma GPT não incorporada em um equipamento, é o advogado por Rosenberg (1998) que utiliza o conceito de GPT, proposto por Bresnahan e Trajtenberg (1995), para examinar a disciplina de engenharia química como uma tecnologia desse tipo e sua relação com o aumento dos setores de aplicação – em especial, petróleo e petroquímicos. O objetivo é explorar a noção de que é útil conceber um conjunto de ideias como constituindo uma GPT. Partindo da hipótese de que a mudança tecnológica é um processo de aprendizagem e que, doravante, toda tecnologia incorpora um grupo de conhecimento, defende que a engenharia

32 Tradução do original em inglês.

química é uma disciplina que incorpora um grupo de conhecimentos, relativos ao *design* de unidades de produção do setor químico, que permite o avanço tecnológico em vários outros setores.

Bekar, Carlaw e Lipsey (2018), a partir de uma análise mais micro das GPTs, definem seis características que servem para identificar uma tecnologia como uma GPT, a saber:

1. Complementaridades com um cluster de tecnologias que o definem e dão suporte: na medida em que a GPT evolui surgem complementaridades tecnológicas multidirecionais dentro do crescente conjunto de tecnologias que os definem e dão suporte.
2. Complementaridades com um cluster de tecnologias que ele habilita: a GPT permite inúmeras novas invenções e inovações, nos mais variados e numerosos setores da economia, muitas das quais tecnicamente impossíveis ou economicamente inviáveis sem a GPT e que não são identificáveis como parte da GPT (ou seja, com o cluster de tecnologias que define e dá suporte a GPT). O que acaba por influenciar a evolução da própria GPT.
3. Complementaridades com um cluster de tecnologias que normalmente incluem aquelas que são social, política e economicamente transformadoras: na medida em que a GPT expande o alcance e a variedade de suas aplicações, as tecnologias envolvidas têm impactos significativos na estrutura econômica e, muitas vezes, nas estruturas políticas e sociais, exigindo mudanças ou abrindo oportunidades para inovações nas tecnologias que são ou serão incorporados nessas estruturas.
4. Não há substitutos próximos: as GPTs criam complementaridades com muitas ou a maioria de suas aplicações que são: (i) únicos - nenhuma outra combinação de tecnologias pode produzir a aplicação - e (ii) Leontief - sem ele, todo o sistema não funcionaria.
5. Tenha uma ampla variedade de aplicações: a ampla gama de complementaridades da GPT (pontos 1 a 3) resulta em GPTs com vários usos ou um único uso genérico que possui muitas aplicações econômicas em grande parte ou em toda a economia.
6. Inicialmente bruto, mas evoluindo em complexidade: as tecnologias que evoluem para se tornarem GPTs geralmente começam de forma bruta e incompleta, geralmente com uma gama restrita de usos. Isso faz com que a identificação de qualquer GPT dependa de um julgamento sobre quando a tecnologia desenvolveu complementaridades tecnológicas suficientes e que agora possui todas as outras cinco características nesta

lista. Podemos saber quando uma tecnologia claramente não é uma GPT e quando claramente é uma, mas a mudança de uma para a outra é um processo contínuo, sem um ponto claro no tempo.

A partir das características definidoras, Bekar, Carlaw e Lipsey (2018) propõe uma definição de GPT mais inclusiva para tecnologias de usos múltiplos e de uso único (amplamente usada versus ter muitos usos): uma GPT é uma tecnologia única, ou um grupo de tecnologias intimamente relacionadas, que possui muitos usos (múltiplos usos), ou um único uso genérico desde que seja amplamente usado, na maior parte da economia. Também é tecnologicamente dinâmica no sentido em que evolui em eficiência e alcance de uso por si só, e é complementar com muitos setores à jusante onde estes usos possibilitam cascatas adicionais de invenções e inovações.

Bekar, Carlaw e Lipsey (2018) argumentam que a definição de GPT que propõe, bem como a lista das seis características definidoras, não inclui os possíveis efeitos macroeconômicos dessa tecnologia, como transformações econômicas, sociais e políticas mais amplas que frequentemente acompanham sua evolução. Esse aspecto, em geral, distingue as várias definições de GPT encontradas na literatura e está presente na definição inicial de Bresnahan e Trajtenberg (1995) apresentada anteriormente. No entanto, os autores apontam a existência de “tecnologias GPTs que são usadas em toda a economia, que possibilitaram muitos desenvolvimentos que não poderiam ter ocorrido sem eles, mas que não tiveram os efeitos macro perturbadores” (BEKAR, CARLAW E LIPSEY, p.1013) como aqueles encontrados quando da introdução da maioria das novas GPTs. O que poderia ser explicado pelo fato de que “se encaixam bem na estrutura atual da economia - o que Lipsey et al. (2005) chamam de estrutura facilitadora - e, portanto, ao contrário da maioria dos outros GPTs, eles não exigem grandes mudanças estruturais e, portanto, não induzem prolongados impactos agregados para se tornar efetivo” (p.1013). O reconhecimento da existência desse tipo de GPT deriva diretamente das características definidoras estabelecidas pelos autores e servem ao propósito de enriquecer a discussão sobre GPTs, principalmente no que tange a sobreposição de múltiplas GPTs, cuja análise da interação possui grande potencial para enriquecer o debate. Assim como uma definição de GPT mais inclusiva permite um conjunto maior de tecnologias desse tipo para estudo. Principalmente o estudo dos fatores do que torna uma GPT uma revolução tecnológica e seu respectivo paradigma ou a gênese de um paradigma tecnológico. Primeiro um olhar micro

para enxergar características e elementos definidores da dinâmica e depois extrapolar para efeitos macro.

Outra questão importante quando da identificação de uma GPT, salientada por Bresnahan (2010) e Bekar, Carlaw e Lipsey (2018), refere-se ao risco de pensar a GPT como sendo incorporada em um único artefato genérico ao invés de conceber GPT como um sistema de tecnologias em evolução. Defendem que uma GPT não pode ser identificada por nenhum artefato que foi inovado em um determinado momento, porque o conhecimento da GPT evolui ao longo do tempo à medida que sua própria eficiência e variedade de aplicativos mudam. Como raramente existe um único artefato tecnológico que possa ser apontado como um GPT, essas tecnologias constituem um conjunto de conhecimentos em evolução que quase sempre é composto por muitos artefatos cooperantes. Quando as GPTs são chamadas de tecnologia única (como uma abreviação para o processo evolutivo de como fazer algo), pode ser prejudicial dado que oculta a longa evolução e a lenta acumulação de conhecimento tecnológico que é a história típica de uma GPT. Uma implicação da definição de uma GPT como um conjunto de conhecimentos em evolução é que raramente existe uma data definitiva para o surgimento de uma GPT ou seu desaparecimento

Bresnahan (2010) também adverte sobre “evitar pensar demais na importância de qualquer tecnologia específica³³” (p.772) quando discorre sobre a recombinação entre inovações complementares existentes (tecnologias anteriores) com uma GPT emergente: “em vez disso, é preciso dizer que o que gera progresso econômico real é o cluster de uma GPT e inovações em aplicações, que juntas unem a oportunidade técnica e a oportunidade do mercado” (p.772)³⁴.

A distinção feita por Lipsey et al. (2005) sobre o que chamaram de *general purpose principles* (GPPs) e *general purpose technologies* (GPTs) se mostra útil na tarefa de identificação de uma GPT: a diferença entre GPPs e GPTs é que tecnologias (GPTs) representam conhecimento subjacente a alguma atividade específica criadora de valor (ou seja, o conhecimento de como produzir e usar algo) enquanto princípios (GPPs) são usados para desenvolver uma gama de tecnologias. Por exemplo, motor a vapor Corliss era uma GPT inspirada na idéia do motor a vapor como um GPP; a ideia de eletricidade é uma GPP, o conhecimento de como gerar,

33 Tradução do original em inglês.

34 Tradução do original em inglês.

distribuir e utilizar (incorporado nos sistemas de produção e distribuição e nos diversos aparatos alimentados por esse tipo de tecnologia) é uma GPT. Uma GPP se transforma em uma GPT por muitas invenções que a tornam uma tecnologia prática. Segundo Bresnahan (2010) é um erro, do ponto de vista econômico, concentrar-se em uma tecnologia específica e restrita na análise de GPTs. Em vez disso, o foco deve estar em um princípio de uso geral como um dos atributos de uma GPT. No caso de interesse, o da WWW, fica claro que o princípio geral (o GPP) é a ideia de conectividade e o conhecimento de como gerar/produzir e usar está incorporado na criação da WWW (conhecimento de criação de redes que é uma das muitas formas possíveis dentro do contexto maior da Internet).

Vale ainda fazer alguns comentários sobre o processo de difusão da GPT onde, normalmente, se observa um “*delay*”, ou seja, a existência de um intervalo entre a data da invenção e a data da realização plena dos ganhos econômicos. Assim, a difusão da GPT é uma fonte de intervalos entre a invenção bruta, ou inicial, e o crescimento de produtividade e econômico final. O padrão que se apresenta é o de uma difusão lenta no início e de posterior aceleração, podendo ocorrer diferentes taxas de difusão para a mesma tecnologia em diferentes aplicações, dependendo das condições de mercado. Dentre as possíveis razões para esse padrão “*slow-rapid*” de difusão, ou para formato clássico de S da curva de difusão, tem-se as restrições de oferta que abrangem as estruturas de custo e as características da tecnologia. Assim como, também, as restrições de demanda: demandantes heterogêneos, usuários com menor renda, custo de ajustamento de aprender com a adoção. E, não menos importante, a existência de redes comunicando informações sobre novas tecnologias com diferentes velocidades ao longo do tempo (redes com formato simétrico, em geral, apresentam “*S-shaped diffusion*” curve da informação. A forma de difusão - velocidade - irá depender do formato da rede).

Além das causas tradicionais, acima citadas, para o formato S da curva de difusão, a estrutura básica da GPT acrescenta mais uma razão para esse formato com as continuadas melhorias na GPT e os mecanismos de feedback entre as inovações GPT e dos setores de aplicação. A parte mais íngreme da curva, ou o período de maior velocidade de adoção da nova tecnologia, surge quando o ciclo de feedback da inovação decola. E tal ciclo é forte quando existem muitos setores de aplicação adotando a nova tecnologia, o que pode incluir todos os setores da economia.

Em resumo, conforme apontou Albuquerque (2019), a importância das inovações radicais, ou de GPTs, na dinâmica tecnológica e econômica de longo prazo tem sido reforçada ao longo do

tempo. Ademais, observa-se uma pluralidade de fontes potenciais de inovações radicais, indicando que mudanças podem estar ocorrendo no antigo sequenciamento de revoluções tecnológicas, pois agora elas poderiam começar a surgir simultaneamente. O que estaria diretamente relacionado à ocorrência de GPTs temporalmente justapostas, ou à superposição de múltiplas GPTs, pois países e setores diferentes, na origem de inovações radicais, podem vir a se desenvolver. Muitas destas GPTs se sobressaem, com impacto suficiente para gerar perturbações importantes na dinâmica do sistema e a elaboração teórica permite captar essa multiplicidade. O resultado mais geral da contribuição teórica dessa elaboração é oferecer um cenário mais turbulento da dinâmica econômica de longo prazo, com perturbações importantes no sistema mais frequentes do que as inicialmente percebidas por Kondratiev e Schumpeter, (ALBUQUERQUE, 2019).

1.3. Considerações finais.

Em resumo, o arcabouço teórico aqui discutido é resultado da percepção de que a mudança técnica apresenta um padrão relativamente ordenado que deriva da natureza acumulativa do conhecimento tecnológico. Cada corpo particular de conhecimento (ou cada paradigma) molda e restringe as taxas e a direção da mudança técnica. Ademais, a mudança técnica é parcialmente impulsionada por repetidas tentativas de lidar com desequilíbrios tecnológicos que a própria mudança cria, ou seja, os maiores impulsos para a inovação derivam dos desequilíbrios entre as dimensões técnicas que caracterizam uma trajetória. Esta, sendo definida como o processo normal de progresso tecnológico dentro das fronteiras de cada paradigma, (DOSI, ORSENIGO, LABINI, 2005 e 2007). A aprendizagem, dessa forma, é local e cumulativa. Local no sentido de que a exploração e o desenvolvimento de novas técnicas e arquiteturas de produtos provavelmente ocorrerão nas proximidades das técnicas e arquiteturas já em uso. Cumulatividade representa a propriedade de que os desenvolvimentos tecnológicos atuais muitas vezes se baseiam em experiências anteriores de produção e inovação, procedem por meio de sequências de conjunturas específicas de resolução de problemas.

Sobre a emergência e seleção dos paradigmas que são desenvolvidos, sabe-se que a demanda e fatores contextuais (como instituições e dinâmicas sociais) podem determinar a direção da mudança técnica entre as várias possibilidades, no entanto, as fontes dessas possibilidades (dos paradigmas tecnológicos) são os avanços na ciência e tecnologia. Assim, a demanda influencia na seleção entre paradigmas competidores e no curso do desenvolvimento do paradigma

estabelecido através da influência sobre as trajetórias estabelecidas por este. Ademais, mudança no paradigma tecnológico implica mudanças nas trajetórias e, como consequência, com diferentes bases de conhecimento e diferentes protótipos de artefatos, as dimensões tecno-econômicas da inovação também variam.

Quando se assume que a inovação tecnológica é o resultado de um processo de acumulação de conhecimento, ou um processo evolucionário, onde as tecnologias atuais derivam das tecnologias precedentes, sendo possível construir uma linha hereditária entre as diversas tecnologias, ou, que uma tecnologia é constituída pela combinação de outras tecnologias, pode-se analisar a WWW como constituída e possibilitada por tecnologias que já existiam e estavam disponíveis, como os protocolos de comunicação da Internet (a tecnologia de redes), os sistemas de hipertexto, a tecnologia de *packet switched*, de *point and click* (mouse), os computadores pessoais, entre outras, como será descrito no próximo capítulo. No entanto, a tecnologia da WWW, foi exclusivamente capaz de colocá-las juntas de forma a proporcionar um acesso fácil e intuitivo para todos os usuários que desejassem navegar na Internet, sendo, assim, popularizada. Como sugeriu Arthur (2009), a tecnologia pode ser pensada tanto a partir de uma visão mais abstrata quanto de uma específica. Dessa forma, analisando a Web, tem-se a *World Wide Web* como uma rede de computadores específica dentro da Internet, materializada por navegadores com seu protocolo http como a tecnologia específica em análise. No entanto, abstratamente, tem-se a ideia de conectividade universal que essa rede proporcionou, com a popularização do uso da Internet e, em especial, do uso comercial desta, que é o princípio geral desta tecnologia.

Essa percepção sugere a existência de um paradigma tecnológico construído em torno de um princípio base que seria a conectividade. Ou ainda, mais especificamente, a ideia de que, se pensado enquanto um domínio tecnológico, no sentido apresentado por Arthur (2009), o paradigma da conectividade emerge e cristaliza-se ao redor da WWW, que seria uma “*novel enabling technology*”, ou uma GPT. E, portanto, tal paradigma teria crescido de forma orgânica a partir desta GPT. Em geral, como argumentado por Bresnahan (2010) e Bekar, Carlaw e Lipsey (2018), uma *General Purpose Technology* representa uma dinâmica inovativa composta por um sistema complexo de interações entre seus numerosos componentes³⁵. Assim, o presente

35 A literatura recente sobre redes tem argumentado que sistemas de inovação são sistemas complexos e, portanto, passíveis de serem representados e analisados pela metodologia de redes.

trabalho, objetivando o estudo desse sistema complexo iniciado pela GPT da WWW, utilizará a abordagem e metodologia analítica de redes para investigação de suas características e dinâmicas, principalmente a partir de seus aspectos científicos e tecnológicos. Isto será demonstrado nos capítulos 3 e 4 quando da descrição da criação das redes e sua subsequente análise.

CAPÍTULO 2: A WORLD WIDE WEB

O presente capítulo se ocupará da história da *World Wide Web*, considerando o período que a antecede, seu processo inovativo, bem como seus primeiros momentos de difusão. A WWW surge a partir da disponibilidade e combinação de outras tecnologias, a saber: a Internet (networks), o sistema de hipertexto, o mouse, os computadores pessoais etc. Sendo assim, a primeira seção do atual capítulo apresentará uma descrição dessas tecnologias antecedentes a partir de uma perspectiva histórica, contextualizando o surgimento da WWW. Posteriormente, a segunda seção se debruçará sobre o processo inovativo da WWW e sua subsequente difusão e desenvolvimento, abordando os primeiros momentos do uso comercial da Internet, possibilitada pela tecnologia da WWW, com a criação e entrada de empresas e novos setores de atividade.

2.1. A Internet: contextualizando.

Segundo Gillies e Cailliau (2000) e Greenstein (2015), a Web é caracterizada por uma difusão generalizada e de singular sucesso que, com frequência, é usada como sinônimo para a Internet. A Internet inaugura uma nova área tecnológica, de redes de computadores – *networking* –, e pode ser entendida como sendo uma rede mundial de “estradas” eletrônicas de informação. A Web, por sua vez, constitui um dos muitos serviços a utilizá-la, provavelmente a mais popular e cujo surgimento potencializou o uso da Internet. Constituindo-se em uma rede dentro de outra rede. Vale, então, uma diferenciação entre as duas tecnologias e uma descrição mais detalhada da história da invenção e difusão de ambas.

A ideia da Internet surge em meados dos anos 1960, em um contexto de guerra fria caracterizada pela ameaça de guerra nuclear, preocupações com armas nucleares e corrida espacial. Foi motivada pela preocupação de como manter a comunicação segura e incorruptível em casos de ameaças militares, por um lado, e com a ineficiência dos sistemas de *circuit-switched* para a comunicação entre computadores, por outro lado. Com o objetivo de criar uma rede para conectar pesquisadores da defesa e campos militares entre si, inicialmente se concentrou numa rede com pouco mais de doze computadores, a ARPANET, que se localizavam em campus universitários, em sua maioria. Apesar de, inicialmente, ser um projeto de inspiração militar, o envolvimento e parceria do meio acadêmico foi preponderante e fundamental, principalmente dos departamentos de ciências da computação, que eram onde os computadores se encontravam

(grandes *mainframes* da IBM, por exemplo), numa época em que os computadores eram bem distintos dos desktops hoje disponíveis. A Internet, então, inaugura a tecnologia de redes, colocando os computadores no centro dos novos meios de comunicações, (ABBATE, 1999; YOFFIE, 1997; GILLIES E CAILLIAU, 2000).

Os Estados Unidos, em um surto de investimento em ciência e tecnologia, criam, nos idos de 1960, a *Advanced Research Projects Agency* (ARPA) para pesquisas militares, motivados pelo pioneirismo russo no lançamento do satélite Sputnik I, no final da década de 1950, que evidenciou o gap tecnológico entre as nações, mudando o pensamento sobre pesquisa nos EUA, (HAFNER E LYON, 1998; ABBATE, 1999, GILLIES E CAILLIAU, 2000). Esta agência, por sua vez, continha um núcleo de pesquisa sobre comunicações, o *Information Processing Techniques Office* (IPTO), inicialmente sobre o comando de J. C. R. Licklider, que era responsável pelo setor de comando e controle cujo objetivo era aplicar todo o potencial dos computadores em situações de guerra e batalha. No entanto, enxergando um maior potencial nos computadores, para além de somente fazer contas rápidas e avançadas, muda dessa ênfase militar para uma pesquisa mais ampla sobre técnicas computacionais avançadas, com especial ênfase em como criar uma rede de computadores, criando o IPTO, por onde passaram os grandes pioneiros da Internet.

Uma primeira versão do que viria a ser a primeira rede de computadores, a ARPANET, foi projetada por Larry Roberts, propondo conectar computadores ao redor do país através de ligações telefônicas entre eles, usando, assim, a rede telefônica já existente para a comunicação. A dificuldade a ser superada nesse primeiro plano residia no fato de que, como os computadores possuíam distintos sistemas operacionais, para cada ligação que um determinado computador desejasse fazer, seria necessário aprender a linguagem de comunicação específica do destinatário. A superação desse obstáculo foi proposta por Wesley Clark e consistia no uso de pequenos computadores como roteadores de dados ao longo da rede, assim os grandes computadores da rede precisariam aprender apenas uma única linguagem de comunicação, a dos roteadores. Roberts acrescentou essa solução ao design da ARPANET denominando esses computadores roteadores de *Interface Message Processors* (IMPs). Assim, cada instituição que fosse conectada à rede, ou cada novo nó na rede, receberia um IMP. Uma característica que marca os primeiros momentos da Internet reside na dificuldade e especificidade do conhecimento para se comunicar na rede, justificando a circunscrição da rede aos usuários acadêmicos por duas décadas, antes do surgimento da WWW e sua popularização do uso da

Internet via facilitação do uso, ou seja, permitindo um uso mais intuitivo. Outro componente inovador na criação da rede foi a escolha da tecnologia de *packet switching* como meio de comunicação da ARPANET que, vale lembrar, foi inicialmente concebida para conectar os maiores computadores do país, os quais eram grandes e dispendiosos e existiam em menor escala em universidades e instituições de pesquisa.

A Internet é uma coleção de redes de computadores comunicando entre si através do que se denominou “*packet switching*”, tecnologia responsável pela divisão da informação em vários pedaços – ou pacotes - do mesmo tamanho para envio através da rede. Assim, todas as comunicações entre computadores na Internet acontecem através de protocolos (a linguagem de comunicação) e da divisão da informação em pequenos pacotes enviados através de estações de roteadores eletrônicos aos seus destinatários (GILLIES e CAILLIAU, 2000). Existem muitas redes de *packet switching* diferentes, operando protocolos específicos e sendo conectadas pela Internet. Ou seja, existem protocolos especializados governando serviços específicos como e-mail, transferência de arquivos, a WWW, LAN (Local Area Network), entre outras. A tecnologia de *packet switching* foi idealizada por dois cientistas separadamente: Paul Baran idealizou uma rede a prova de bombas onde os pacotes de informações pegariam rotas alternativas caso alguma rota fosse bombardeada; e Donald Davies, preocupado com a insuficiência do sistema usado para ligações telefônicas, *circuit switched*, para comunicação entre computadores.

Assim, o primeiro projeto do que hoje conhecemos como a Internet, a ARPANET, sobre o comando de Larry Roberts, contou com a participação de Donald Davies que transferiu o conhecimento da tecnologia de *packet switching*. É importante ressaltar que a tecnologia de *packet switching*, tida como muito avançada e complexa, enfrentou grande resistência da comunidade de comunicações a época (ABBATE, 1999; GILLIES E CAILLIAU, 2000). Abbate (1999) aponta que, inicialmente, os desenvolvimentos da rede incorporaram princípios militares como a capacidade de sobrevivência e grande capacidade de comunicação, ao invés de princípios mercadológicos. Em 1967, a ARPA realiza um contrato com Elmer Shapiro para invenção do design da rede de IMPs e, no ano seguinte, lança um edital para o fornecimento dos primeiros quatro IMPs, que foram estabelecidos em instituições de P&D, em computadores onde o conhecimento técnico e específicos necessários para fazer a rede funcionar estava disponível. O primeiro nó da ARPANET foi estabelecido na University of California at Los Angeles (UCLA), em 1969, no laboratório de Leonard Kleinrock, tendo um Honeywell DDP-

516³⁶ como o primeiro protótipo de IMP. Alguns meses depois a Stanford Research Institute instala o segundo IMP e a primeira rede de *packet switching* se estabelece. A primeira mensagem enviada na rede foi LO e o sistema falhou após isso. Os terceiro e quarto nós da rede, ou IMPs, foram instalados, respectivamente, na University of California at Santa Barbara e Utah. Completando, assim, a rede original composta por quatro hosts. Até o final de 1970 a rede crescia a uma taxa de um host instalado por mês (GILLIES E CAILLIAU, 2000). Os primeiros protocolos da ARPANET, que permitiam a comunicação na rede de IMPs, foram telnet e *file transfer protocol* (FTP)³⁷. O primeiro permitia que um computador realizasse o login remoto em outro computador e o segundo, que arquivos fossem transferidos entre eles. Estes protocolos foram substituídos pelo Network Control Protocol (NCP), um protocolo mais geral e cujo trabalho era apenas entregar os pacotes de informação sem discriminar o seu tipo.

Simultaneamente a criação e desenvolvimento da ARPANET, e da difusão do conhecimento em networking, outros esforços de criação de redes de computadores também aconteciam em outros países, como a rede desenvolvida por Donald Davies no National Physical Laboratory, na Inglaterra, que careceu de financiamento para avançar. Tem-se também a ALOHANET, desenvolvida pela University of Hawaii com o objetivo de conectar as estações de pesquisa remotas, localizadas nas diversas faculdades ao longo de quatro ilhas, com o computador central IBM. A ALOHANET foi a primeira *packet radio network* e atraiu a atenção da ARPA pelo potencial de aplicação militar de uma rede que não utilizava cabos e que, portanto, poderia ter nós móveis, o que se denominou *Packet broadcasting*. Mais tarde foi conectada a outras redes por satélite, lançando as bases para terminais móveis de Internet, o que lançou as bases para o surgimento dos celulares com Internet. A rede francesa Cyclades, desenvolvida com incentivo do governo e que tentou uma abordagem diferente da ARPANET, menos baseada no hardware e mais no software, também construída com forte presença do setor acadêmico, que possuíam os primeiros nós. No entanto, uma mudança do grupo político no poder inviabilizou a continuação do projeto.

36 Importante notar a ocorrência de desenvolvimento concomitante da rede (networking) e dos computadores ao longo do tempo. O desenvolvimento da rede encontrou muitos obstáculos no nível de desenvolvimento dos computadores.

37 Esses primeiros protocolos foram desenvolvidos pelo Network Working Group composto pelos programadores dos primeiros quatro institutos que hospedaram os IMPs.

Em 1983, a ARPANET adota um novo conjunto de protocolos e é considerado o ano em que a Internet, como a conhecemos, se inicia. Esse conjunto básico de protocolos em funcionamento na Internet é chamado TCP/IP. O IP, *Internet Protocol*, é a língua universal de comunicação entre computadores, é o protocolo responsável por carregar os pacotes de informação. O TCP, *Transfer Control Protocol*, opera na divisão da informação em pacotes, sua correta identificação e remontagem quando os pacotes chegam a seu destino. Vale notar que redes individuais podem usar, internamente, protocolo diferentes, porém, quando seus respectivos *gateway* em operação as conectam à Internet, devem carregar pacotes usando o protocolo TCP/IP. Este conjunto de protocolos eram considerados mais simples e de fácil utilização, contribuindo para a difusão da Internet. Outro fator que colaborou com a evolução da Internet foi a criação, em meados dos anos 1980, do *Domain Name System* (DNS) que criava regras de endereçamento para cada nó na rede e que, mais tarde, foi largamente utilizado na WWW.

Segundo Gillies e Cailliau (2000), Abbate (1999) e Hafner e Lyon (1998), uma das características mais importantes da Internet reside na forma como ela foi desenvolvida, sendo o resultado da colaboração de muitos cientistas ao redor do mundo, não pertencendo a uma companhia e/ ou país específico. O objetivo principal dessa “sociedade” era facilitar a interação homem-computador, a partir do desenvolvimento do campo emergente de redes, garantindo uma evolução da tecnologia como se esta possuísse “vida própria”. No entanto, na medida em que evolui e difunde, atrai o interesse de governos e companhias de telecomunicação, que antes se opuseram a tecnologia de *packet switching* por seu alto nível de incerteza, trazendo uma tentativa de organização e regulação da rede, bem como de exploração econômica. Não obstante, como salientam os autores, a natureza anárquica de seu desenvolvimento prevalece: “[...] quem governa a Internet? Onde fica a sede? Quem é responsável? E a resposta é simplesmente: consenso aproximado e código em execução [...]” (GILLIES E CAILLIAU, 2000, p.65)³⁸. Tal característica teria garantido a prevalência da Internet sobre outros padrões de rede em curso (como o protocolo *Open Systems Interconnection* da *International Standard Organization*, por exemplo), que, em geral, tentaram impor um design de rede que seria fácil taxar, em detrimento da facilidade do uso, ou seja, não priorizava as questões técnicas. E, portanto, foram abandonadas. A ARPANET não foi construída para propósitos comerciais. Outros fatores que contribuíram para o predomínio dos protocolos TCP/IP foram a adoção da Ethernet como a *Local Area Network* mais utilizada e difundida que já incorporava a linguagem

38 Tradução do original em inglês.

TCP/IP; a adoção pela comunidade acadêmica do sistema operacional UNIX que também já incorporava tais protocolos; e o surgimento de outros projetos de *network* como o SUPERJANET que adotava o TCP/IP, estabelecendo assim a Internet e fazendo com que as demais redes fizessem a migração para esse conjunto de protocolos.

Paralelamente ao desenvolvimento da ARPANET, que inicialmente se circunscrevia à comunidade de pesquisa em computação, outras redes foram surgindo, como a BITNET, também voltada para a comunidade acadêmica e resultado da conexão entre computadores IBM que possuíam um software de network incluso, o *Remote Spooling Communications Subsystem* (RSCS), uma rede composta por universidades nos EUA e que ganhou links na Europa constituindo a rede denominada EARN. Essa rede foi pioneira no envio de email's e no uso de fóruns de discussão, usando a tecnologia LISTSERV. Já a comunidade da física de partículas³⁹, que escolheu adotar os computadores DEC, compartilhavam a rede DECNET, também possibilitada pelo software já incluso nos dispositivos. Ambas as redes usavam, inicialmente, os protocolos X25 da ISO, no entanto, estes se mostraram inadequados para aplicações que demandavam mais que a realização de login remoto e envio de email's, sendo substituídos pelo TCP/IP que era de uso fácil e gratuito. Outras redes maiores, conhecidas como rede de redes ou Internetworking, também surgiram, como a CSNET, criada pela *National Science Foundation* (NSF), que tinha a ARPANET como um de seus componentes, e a NSFNET, também da NSF. Nota-se, assim, uma difusão das redes e o consequente compartilhamento de informações.

É importante ressaltar o fato de que os protocolos governando serviços na Internet foram, por duas décadas, dominados pela comunidade acadêmica, já que se exigia conhecimento especializado e complexos comandos para acessar informações na rede e, naturalmente, para inseri-las também, além do fato de ter sido a comunidade acadêmica uma grande parceira do setor militar no desenvolvimento da rede. Ademais, era natural que a Internet surgisse e crescesse dentro do meio acadêmico dado que era lá onde estavam os computadores, pois não havia computadores pessoais e muito menos a demanda por eles⁴⁰. A difusão de Internet, de

39 Gillies e Cailliau (2000) e Berners-Lee e Fischetti (2000) salientam a relação direta e indutiva dos desenvolvimentos na área da física de partículas sobre as áreas de ciência de networking e computação. Muitos spillovers ocorreram diante da necessidade de computadores mais potentes e velozes, bem como de comunicação entre cientistas dos mais diversos países para o andamento das pesquisas.

40 Gillies e Cailliau (2000) apontam que, para a maioria da população, era impensável ter um computador em casa e na maioria dos empregos, dados os custos e, até mesmo, a ausência de propósito. Ou seja, a maioria das pessoas não concebia o motivo de se usar um computador em suas atividades rotineiras.

modo significativo, se deu a partir da revolução iniciada pelos computadores pessoais que, por sua vez, foi possibilitada pelo surgimento da tecnologia do microprocessador (resultado do transistor e dos circuitos integrados). Este, permitiu a criação de computadores menores, mais potentes e mais velozes. Tem-se, então, a explosão de companhias fabricantes de *personal computers* (PC), como a Xerox com seu Alto, a Apple com seu Macintosh, IBM com IBM PC, entre tantos outros. Além de companhias produzindo os mais variados softwares, como os sistemas operacionais, por exemplo. A popularização dos computadores pessoais, aumentando a demanda por acesso e conexão, e o surgimento da WWW⁴¹ com sua simplicidade de *point and click*, propiciou o acesso à Internet a qualquer indivíduo que possuísse um computador e uma linha de telefone. Durante alguns anos, a maioria das pessoas acessavam a Internet através de uma chamada telefônica para um Provedor de Serviço de Internet (ISP), tal método se baseava no uso de modems que faziam a ligação entre a real Internet no ISP e o computador. No entanto, outras tecnologias foram surgindo, como a fibra ótica e o *wireless*, que permitiram o uso cada vez mais veloz da Internet e confirmando o caráter acumulativo e evolutivo da mudança tecnológica:

Acima de tudo, a própria noção do que é a Internet - sua estrutura, seus usos e seu valor - mudou radicalmente ao longo de sua existência. A rede não era originalmente um meio de comunicação interpessoal; a intenção era permitir que os cientistas superassem as dificuldades de rodar programas em computadores remotos. A Internet atualmente dirigida comercialmente e orientada para a comunicação surgiu somente após um longo processo de reestruturação técnica, organizacional e política. (ABBATE, 1999, p.2)

Vale ainda uma nota sobre um personagem e sua teoria de preponderante importância quando se pensa os campos das telecomunicações e, em especial, da informação, no geral. Claude Shannon e sua teoria matemática da informação é considerado o precursor dos avanços na comunicação de informações. Debruçou-se sobre as questões de como estabelecer comunicações de forma eficiente, como entender e definir o que é a informação de fato e, em última instância, como representá-la da menor forma possível para que sua transmissão contasse com a menor distorção, ou ruído, possível. Usando a teoria binária para definir informação, publicou em 1948⁴² o artigo denominado “*Uma teoria matemática da comunicação*” onde lança as bases para a comunicação de informações através de circuitos elétricos. Foi o criador

41 O surgimento da WWW também é um resultado direto da difusão da Internet, dos computadores pessoais e dos sistemas de hipertexto.

42 Mesmo ano em que a Bell Telephone Laboratories (Bell Labs) anunciou a invenção de um pequeno semicondutor eletrônico, o transistor, o catalizador da revolução na eletrônica.

do *bit*, que é a menor unidade de informação reconhecida, ou unidade fundamental da informação, conformando uma nova unidade de medida. Permitindo que toda e qualquer informação pudesse ser representada por uma sequência de bits. É considerado o precursor das tecnologias de computação (através do uso da lógica Booleana), e, por muitos, da era digital. Como, também, é a ele atribuído a criação de um novo campo da ciência: a da informação, (SONI E GOODMAN, 2017; GLEICK, 2011;).

2.2. *World Wide Web*: antecedentes, criação e desenvolvimentos posteriores⁴³.

Essa seção se destina à uma descrição histórica do surgimento da *World Wide Web* (WWW), compreendendo uma apresentação do contexto tecnológico de sua invenção, seus antecedentes, o decorrer do processo inovativo dos primeiros momentos de difusão.

Michael Dertouzos escreveu em “*Weaving the Web*”, de Berners-Lee e Fischetti (2000), que uma história se destacava dentre as milhares de informações sobre a *World Wide Web* (WWW): “a da criação e evolução contínua desta coisa nova incrível que está surgindo para abranger o mundo e se tornar uma parte importante e permanente de nossa história”, (p. vii)⁴⁴. E, ainda, “uma abordagem única que altera o curso da humanidade” (p.ix)⁴⁵. O interessante dessas citações é notar a afirmação do poder transformador dessa tecnologia ainda tão próximo da sua criação. Dertouzos (2000) ainda segue dizendo que “a inovação de Tim também é única” na medida em que possibilitou a criação de um enorme mercado de informações, onde informação e serviços de informação eram vendidos, comprados ou trocados livremente entre indivíduos e organizações.

A imprensa, o rádio e a televisão nunca chegaram perto; tudo o que podem fazer é espalhar as mesmas informações de uma fonte para vários destinos. Nem a carta ou o telefone podem se aproximar do poder da Web, porque embora essas mídias possibilitem trocas individuais, elas são lentas e desprovidas da capacidade do computador de exibir, pesquisar, automatizar e mediar. Surpreendentemente - em comparação com a imprensa de Gutenberg, o telefone de Bell e o rádio de Marconi - e muito antes de atingir sua forma final, a Web de Berners-Lee já estabeleceu sua singularidade (DERTOUZOS, 2000, p.vii).⁴⁶

43 Essa seção está amplamente baseada nos livros de Berners-Lee e Fischetti (2000) e Gillies e Cailliau (2000).

44 Tradução do original em inglês.

45 Tradução do original em inglês.

46 Tradução do original em inglês.

Dertouzos (2000) e Yoffie (1997) salientam, ainda, que cientistas estiveram por anos fixados em dois elementos, o hipertexto e a redes de computadores, mas, apenas Berners-Lee concebeu como conectá-los e criar a Web. “Acontece que eu apareci com o tempo, e com o interesse e a inclinação certos, depois que o hipertexto e a Internet amadureceram. A tarefa que me restou foi casar os dois juntos” (BERNERS-LEE E FISCHETTI, 2000, p.6)⁴⁷. Dertouzos (2000) pontua também que seu inventor, já nos anos 2000, se preocupava em definir a Web do futuro: enquanto o mundo se ocupava do comércio eletrônico, Berners-Lee já idealizava a Web como um meio que codificaria, através de seus links de informação enormemente distribuídos, o conhecimento e o entendimento humano. Claramente uma antecipação do que se entende atualmente como Inteligência Artificial.

Gillies e Cailliau (2000) comentam sobre a combinação inovadora da Web:

Não é uma surpresa que tão poucas pessoas compreenderam completamente o sonho de Tim. Hipertexto e networks não eram ideias novas. Ambas existiam desde 1960, mas ninguém as havia juntado para criar um sistema de hipertexto global até então. Agora parece difícil acreditar que os benefícios não tenham sido óbvios, porém não eram. Muitas pessoas brilhantes haviam trabalhado nos dois sistemas e colocá-los juntos não havia ocorrido a ninguém [...]. Quando a autoestrada da informação era pouco mais do que uma viela, Tim Berners-Lee havia inventado o equivalente, em networking, a um carro de corrida de Fórmula 1, enquanto todo mundo ainda vagava por aí em Austin Sevens. Não admira que as pessoas não tenham entendido (p.201)⁴⁸.

2.2.1. Antecedentes

A WWW foi inventada dentro da instituição internacional *Conseil Européen pour la Recherche Nucléaire* (CERN)⁴⁹, e é o resultado de sucessivas tentativas, por parte de Tim Berners-Lee – um físico que atuou como consultor na instituição por alguns anos – de dar forma a uma visão que envolvia um crescimento orgânico e descentralizado de ideias, tecnologia e sociedade. Segundo Berners-Lee, “a visão que tenho para a Web é sobre qualquer coisa estar potencialmente conectada a qualquer coisa” (BERNERS-LEE E FISCHETTI, 2000, p.1)⁵⁰. Além de ser motivada por uma antiga inquietação sobre a possibilidade de computadores

47 Tradução do original em inglês.

48 Tradução do original em inglês.

49 Gillies e Cailliau (2000) argumentam que a curiosidade é razão da existência da CERN, sendo o conhecimento seu principal produto. E que as tecnologias requeridas pelas pesquisas realizadas no laboratório, voltadas para física de partículas, geraram inúmeros transbordamentos, dentre os quais a WWW.

50 Tradução do original em inglês.

performarem intuitivamente imitando a mente humana em sua habilidade de completar conexões, Berners-Lee e Fischetti (2000, p. 3)⁵¹: “[...] os computadores eram bons em organização e processamento lógicos, mas não em associações aleatórias. Um computador normalmente mantém as informações em hierarquias e matrizes rígidas, enquanto a mente humana tem a habilidade especial de vincular bits aleatórios de dados”.⁵²

Sobre o processo de invenção da WWW, Berner-Lee descreve o que poderia, perfeitamente, ser a descrição de um processo inovativo em geral – como um processo não isolado e de acumulação de conhecimento contínuo, apesar de não linear:

Os jornalistas sempre me perguntaram qual era a ideia crucial, ou qual era o evento singular, que permitiu que a web existisse num dia quando não havia existia no dia anterior. Eles ficam frustrados quando eu digo a eles que não houve um momento "Eureka!". [...] Inventar a World Wide Web envolveu minha crescente percepção de que havia um poder em organizar ideias de uma maneira irrestrita como a da web. E essa consciência veio a mim precisamente por meio desse tipo de processo. A Web surgiu como a resposta a um desafio aberto, através do turbilhão de influências, ideias e realizações de muitos lados, até que, pelos maravilhosos escritórios da mente humana, um novo conceito se formou. Foi um processo de acréscimo, não a resolução linear de um problema bem definido após o outro (BERNERS-LEE E FISCHETTI, 2000, P.3)⁵³.

A primeira aproximação do que veio a se tornar a WWW, tomou forma em um programa de software denominado Enquire⁵⁴ escrito por Berners-Lee em 1980 quando esteve na CERN⁵⁵ para um trabalho de consultoria de software e, a princípio, foi criado para uso pessoal. O laboratório estava em fase de mudança dos sistemas de controle dos aceleradores de partículas⁵⁶, os quais eram o motivo para a vinda de cientistas e demais pesquisadores, oriundos das mais variadas áreas do planeta, dado que experimentos desse tipo eram caros e necessitavam de esforço conjunto. Assim, as diversas equipes rodavam os experimentos e os indivíduos retornavam ao país de origem com os dados a serem analisados. Isso resultava em sistemas de controle cuja construção era resultante de várias contribuições e colaborações de pessoas

51 Tradução do original em inglês.

52 Berners-Lee criticava os sistemas de organização hierárquicos ao dizer que restringiam o crescimento.

53 Tradução do original em inglês.

54 Um diminutivo para “*Enquire Within upon Everything*” que, segundo Berners-Lee relata, se tratava de um título de um livro que poderia ser considerado um portal para um mundo de informações.

55 Mais precisamente no “*European Particle Physics Laboratory*” em Genebra.

56 Especificamente, o projeto se baseava em criar um novo sistema geral de controle para o *Proton Synchrotron Complex* que já operava antes do estabelecimento de computadores para controle e operação dos aceleradores. À medida que computadores foram sendo acoplados ao centro de controle, cada um continha a responsabilidade de uma atividade específica nos aceleradores, carecendo de um sistema de controle geral. O que motivou o projeto que trouxe vários técnicos, entre eles Berners-Lee, ao laboratório.

(cientistas e físicos), pesquisas, artigos e etc. Na tentativa de entender os sistemas, humanos e computacionais, Berners-Lee percebeu que precisaria criar um sistema que organizasse e ligasse as informações que ia recebendo, com objetivo de auxiliar a lembrar as conexões entre as várias pessoas, computadores e projetos no laboratório (BERNERS-LEE E FISCHETTI, 2000; GILLIES E CAILLIAU, 2000).

O *Enquire* é escrito por Berners-Lee no computador usado para o desenvolvimento de software do grupo de consultores alocados no desenvolvimento dos novos sistemas de controle, dado que, naquela época, era raro que cada membro do *staff* da CERN possuíssem um computador em sua estação de trabalho, fazendo com que cada indivíduo se dirigisse ao que chamavam de *terminal room* para desenvolverem seus programas de software. A ideia base do *Enquire* era armazenar informações conectadas entre si: uma nova informação (um nó da rede que se formava) só poderia ser adicionada ao programa se estivesse conectada com outra já existente (outro nó). Ou seja, havia a obrigatoriedade de enfatizar a conexão (ou *link*) entre as informações e, além disso, especificar a natureza dessa relação. O objetivo era manter um ‘rastros’ sobre quem escreveu o quê, em cada sistema, de cada máquina. O *Enquire* armazenava as informações através de associações randômicas, ao contrário das estruturas matriciais e do tipo árvore comumente utilizadas, foi pensado para permitir dois tipos de links: os internos, que se davam entre os nós de uma mesmo arquivo (ou pasta); e os externos, que compreendiam as conexões entre diferentes arquivos (BERNERS-LEE E FISCHETTI, 2000). Assim, uma rede de informações começava a tomar forma. Eventualmente o contrato de consultoria acabou e Berners-Lee ao deixar a CERN deixou o programa com algum membro do laboratório que repassou para outro e *Enquire* original acabou se perdendo. Vale notar que o software rodava no computador do grupo e não estava conectado em qualquer rede ou Internet (a CERN só adotaria a Internet anos depois).

Entre 1983 e 1984, a antes de retornar a CERN para um *fellowship program*, Berners-Lee, de posse de um Compaq, considerado o primeiro computador portátil disponível, ainda explorando a ideia de conectividade, cria o *Tangle*. A intuição por trás dessa nova tentativa era a de que, dado que os computadores armazenavam informações através de sequências de caracteres e que seus significados (ou sentido) estavam em suas conexões, o importante, então, eram as conexões entre caracteres e não os caracteres em si. Assim, toda vez que uma certa sequência de caracteres era recorrente, *Tangle* criaria um nó que representasse tal sequência. A partir daí, toda vez que essa mesma sequência ocorresse, ao invés de repeti-la, era somente criado uma

referência ao nó correspondente (o nó original). Na medida que mais sequências de caracteres foram armazenadas como nós e que mais referências foram sendo feitas, uma série de conexões foram formadas. No entanto o projeto foi abandonado quando, na execução de um teste, o programa entrou em *loop* infinito.

Tim Berners-Lee retorna a CERN em 1984 e encontra uma instituição que havia crescido e estava em processo de construção de um novo acelerador de partículas, maior em tamanho e capacidade. Cientistas do mundo todo se dirigiam ao laboratório e traziam consigo os próprios computadores, sistemas e programas que, nos últimos anos, haviam evoluído e diversificado. Mesmo quando não estavam na instituição, os pesquisadores continuavam a colaborar com as pesquisas em andamento, o que exacerbava a necessidade de conexão. Assim, Berners-Lee escreveu um programa geral de “*remote procedure call*” (RPC) destinado a facilitar a comunicação entre todos os computadores e redes: “com o RPC, um programador pode escrever um programa em um tipo de computador, mas permitir que ele chame os procedimentos em outros computadores, mesmo que eles sejam executados em sistemas operacionais ou linguagens de computador diferentes” (BERNERS-LEE E FISCHETTI, 2000, p.14)⁵⁷. Além do mais, o programa funcionava em qualquer rede disponível.

Alocado no setor de “aquisição e controle de dados”, um grupo dedicado a captura e processamento dos resultados dos diversos experimentos, Berners-Lee começou a recriar o *Enquire*, que dessa vez rodava em seu Compaq, mas também foi escrito para rodar no VAX *minicomputer* produzido pela DEC e amplamente usado na CERN. No entanto, essa nova versão também permaneceu inacabada.

No entanto, Berners-Lee estava convencido da necessidade de algo como o *Enquire* em uma instituição plural como a CERN:

Além de acompanhar as relações entre todas as pessoas, experimentos e máquinas, eu queria acessar diferentes tipos de informações, como documentos técnicos de um pesquisador, os manuais para diferentes módulos de software, atas de reuniões, anotações rabiscadas apressadamente e assim em diante. Além disso, encontrei-me respondendo às mesmas perguntas feitas com frequência, a mim mesmo, por pessoas diferentes. Seria muito mais fácil se todos pudessem apenas ler meu banco de dados (BERNERS-LEE E FISCHETTI, 2000, p.15)⁵⁸.

57 Tradução do original em inglês.

58 Tradução do original em inglês.

Na tentativa de formalizar essa proposta, Berners-Lee viu que sua ideia se encontrava dentro da categoria geral de “*documentation systems*”, ou seja, de software que permitiam o armazenamento e posterior recuperação de informações. O que se mostrava um desafio dada as numerosas tentativas anteriores dentro da CERN que falharam. Principalmente em função da diversidade de sistemas e redes de computadores que levava os desenvolvedores a exigirem a reorganização do trabalho, por parte dos pesquisadores, para servir ao sistema. Daí surgia a necessidade clara de criar um sistema com regras comuns e simples que seriam aceitas por todos. “Isso significava o mais próximo possível de nenhuma regra” (BERNERS-LEE E FISCHETTI, 2000, p.15)⁵⁹.

A partir dessa percepção, Berners-Lee percebe que, ao invés de um obstáculo, a diversidade de sistemas operacionais, computadores e *networks* poderia ser um rico recurso na construção de seu sistema global de informações. Dessa forma, o formato do projeto começa a se delinear. A primeira escolha é o modelo do sistema: hipertexto. E o plano se resumia a uma visão que combinava os links externos do *Enquire* com hipertexto e os esquemas de interconexão desenvolvidos anteriormente para RPC, ou seja, um programa de *Enquire* capaz de realizar links de hipertexto externos. “Novas teias poderiam ser criadas para ligar diferentes computadores, e todos os novos sistemas seriam capazes de entrar em ação e fazer referência a outros. Além disso, qualquer pessoa que estiver navegando pode adicionar instantaneamente um novo nó conectado por um novo link” (BERNERS-LEE E FISCHETTI, 2000, p.16)⁶⁰. Outra propriedade fundamental do sistema era ser completamente descentralizado, permitindo a adição de novos usuários livremente (sem a necessidade de permissão de uso), o que era considerado o requerimento essencial para o crescimento do sistema. Para tanto, bastava que não fosse introduzido uma base de dados central de links, ou seja, um nó central ao qual tudo o mais estivesse conectado. A maioria dos sistemas disponíveis dependiam de um nó central o que acabava por limitar seu crescimento em algum momento.

Sem a introdução de um link central o resultado seria um sistema flexível, onde não haveria links ou nós especiais, ou seja, qualquer nó potencialmente poderia se conectar com qualquer outro nó. Essa flexibilidade era, segundo Berners-Lee, a chave par um sistema universal. “O

59 Tradução do original em inglês.

60 Tradução do original em inglês.

hipertexto seria mais poderoso se pudesse apontar para absolutamente qualquer coisa. Cada nó, documento - seja lá como for chamado - seria fundamentalmente equivalente de alguma forma. Cada um teria um endereço pelo qual poderia ser referenciado. Eles existiriam todos juntos no mesmo espaço - o espaço da informação” (BERNERS-LEE E FISCHETTI, 2000, p.16)⁶¹.

Em 1988, Berners-Lee percebe a necessidade de formular uma proposta para que o desenvolvimento do “*hypertext documentation system*” se tornasse um projeto formal aprovado dentro da CERN. Até então, todos os desenvolvimentos anteriores haviam acontecido por conta própria, em seu tempo livre e como uma forma de investigação pessoal. Algumas definições já estavam claras: o *Enquire* era a forma de conectar documentos e base de dados, o hipertexto seria o formato comum para exibi-los, o desafio se encontrava, então, em como estabelecer a comunicação entre diferentes computadores com diferentes sistemas operacionais. O grande avanço na direção da conectividade global veio com desenvolvimento da Internet nos anos 1970. Essa estrutura geral de comunicações que era capaz de conectar computadores, moldada por Donald Davis, Paul Barran, Vint Cerf e Bob Kahn, entre outros, começava a se difundir. Assim, a Internet, apesar de já ser uma realidade em curso nos EUA possibilitando a conexão entre universidades e laboratórios por todo o território, encontrava resistência na Europa e, conseqüentemente, na CERN. Berners-Lee e Fischetti (2000) e Gillies e Cailliau (2000) relatam a tentativa Europeia na busca de um conjunto de protocolos de rede próprio, ou seja, uma rede internacional criada por eles.

A Internet pode ser definida como uma infraestrutura de comunicação geral que conecta computadores permitindo uma comunicação indireta através de uma rede, superando as limitações das conexões feitas através de cabos dedicados feitas anteriormente, (BERNERS-LEE E FISCHETTI, 2000; GILLIES E CAILLIAU, 2000). Resumidamente, a Internet pode ser pensada como um conjunto de protocolos padronizados, ou seja, de convenções pelas quais dados são enviados entre computadores⁶². Ela já se encontrava em funcionamento a partir da década de 1970, mas seu uso se dava, majoritariamente, entre *experts* da computação dada a

61 Tradução do original em inglês.

62 “The data are transmitted over various carriers, such as telephone lines, cable TV wires, and satellite channels. The data can be text, an e-mail message, a sound, an image, a software program - whatever. When a computer is ready to send its data, it uses special software to break the data into packets that will conform to two Internet protocols that govern how packets will be shipped: IP (Internet Protocol) and TCP (Transmission Control Protocol). The software labels each packet with a unique number. It sends the packet out over the phone or cable wire, and the receiving computer uses its own Internet software to put them back together according to the labels.” (BERNERS-LEE E FISCHETTI, 2000, P. 18)

dificuldade de transmissão de informação que exigia diferentes passos para estabelecer uma conexão, acessar a informação, transferi-la e decodificá-la. O que poderia exigir programas distintos para cada etapa, dada a diversidade de linguagens presentes nos diversos computadores, sistemas operacionais, e software de programas, dificultando um uso mais generalizado e intuitivo. O surgimento do correio eletrônico (e-mail) facilita o envio de mensagens, mas não criava um espaço no qual a informação pudesse permanecer disponível para acesso e referência. A WWW seria um espaço onde a informação pode persistir.

Berners-Lee enxerga na Internet a possibilidade de superar as dificuldades advindas da existência de diferentes sistemas operacionais e redes, construindo uma ponte entre eles através de um programa que fosse de fácil uso e permitisse que cada indivíduo pudesse manter suas particularidades organizacionais e softwares em seu computador, o que permitia o crescimento global da rede. Ademais, os protocolos da Internet (TCP/IP) começavam a ficar disponíveis para sistemas operacionais diversos, dentre eles os mais utilizados pelos cientistas na CERN: VAX/VMS e Unix. Então, em resumo, a ideia era criar um sistema de acesso à informação através de uma rede, baseado no conceito inicial do *Enquire* - um sistema de documentação -, que permitisse aos indivíduos, de uma forma simples, criar representações e navegar através dos links entre seus documentos, o que seria possibilitado pelo uso de hipertexto. Os protocolos de rede seriam TCP/IP e, através do esquema adaptado de endereçamento do RPC, comunicações e compartilhamento de documentos aconteciam, pois este informava aos indivíduos como atribuir um endereço aos seus arquivos e como requisitar arquivo de terceiros.

Em março de 1989 Berners-Lee entrega sua proposta formalizada, denominada “*Information Management: A proposal*”, para membros da direção da CERN, porém não obteve respostas. Berners-Lee e Fischetti (2000) e Gillies e Cailliau (2000) destacam que a falta de interesse da instituição se baseava, em grande parte, no fato de ser um laboratório de física, que encarava a tecnologia da informação como um meio para um fim, com pouco espaço na estrutura a ser designado a um projeto na magnitude do proposto. Em 1990, ainda aguardando um *feedback*, Berners-Lee repagina o projeto e volta a apresentá-lo e, novamente, não obtém sucesso. A introdução de sua proposta delineava sua visão do problema de gerenciamento de informação dentro da instituição:

“O CERN é uma organização maravilhosa. Envolve muitas milhares de pessoas, muitas delas muito criativas, todas trabalhando em prol de objetivos comuns. Embora sejam nominalmente organizadas em uma estrutura de gerenciamento hierárquica,

isso não restringe a maneira como as pessoas se comunicam e compartilham informações, equipamentos e software entre os grupos. A estrutura de trabalho real observada da organização é uma “web” multiplamente conectada, cujas interconexões evoluem com o tempo. Nesse ambiente, uma nova pessoa chegando ou alguém assumindo uma nova tarefa normalmente recebe algumas dicas sobre com quem seria útil conversar. As informações sobre as instalações existentes e como descobri-las percorrem nas fofocas de corredor e boletins informativos ocasionais, e os detalhes sobre o que deve ser feito são divulgados de maneira semelhante. Tudo considerado, o resultado é notavelmente bem-sucedido, apesar de mal-entendidos ocasionais e esforços duplicados” (BERNERS-LEE, 1989, p.3)⁶³.

Apesar da ausência de respostas institucionais, Tim Berners-Lee mantém sua visão, uma visão mais ampla, na qual toda a informação contida em todos os computadores estivesse conectada. Onde todos os bits de informação, em cada computador do planeta, estivessem disponíveis para todos os indivíduos, um espaço único de informação global. “Uma vez que um bit de informação naquele espaço fosse rotulado com um endereço, eu poderia dizer ao meu computador para alcançá-lo. Por ser capaz de fazer referência a qualquer coisa com a mesma facilidade, um computador pode representar associações entre coisas que podem parecer não relacionadas, mas de alguma forma compartilham um relacionamento. Uma teia de informações se formaria” (BERNERS-LEE, 1989, p.4)⁶⁴.

É relevante citar que conceitos similares foram desenvolvidos por outras pessoas, apesar de não implementados, e que formaram a base de conhecimento para o que veio a ser a Web: em um artigo de 1945 intitulado “*As We May Think*”, Vannevar Bush, descreve uma máquina foto-eletromecânica (Memex) que poderia criar e seguir referências cruzadas entre documentos de microfilmes, o principal objetivo desse artefato mecânico era armazenar, recuperar e editar informações endereçando à preocupação de Bush quanto a avalanche de informações que se avolumava principalmente a partir do avanço das pesquisas científicas, além do desejo de replicar artificialmente o modo de operação da mente humana que aprende por associação ao invés de hierarquização da informação, idealizou o que hoje conhecemos como computadores pessoais capazes de links associativos ao invés de hierárquicos; Ted Nelson, em 1965, cria o conceito de hipertexto, um novo formato não sequencial de texto, e idealiza o projeto Xanadu, no qual toda a informação disponível no mundo fosse publicado em hipertexto; e, por último, o espaço de trabalho colaborativo denominado NLS (oN Line System), demonstrado por Doug Engelbart, onde textos podiam ser visualizados por vários usuários simultaneamente e editado

63 Tradução do original em inglês.

64 Tradução do original em inglês.

em colaboração, introduziu o uso do sistema de hipertexto em uma ferramenta para trabalho em grupo⁶⁵, sua motivação era lidar com o aumento crescente da informação, antecipado por Bush, e com a crescente complexidade dos problemas humanos e a diminuição do tempo para lidar com eles, assim uma ferramenta que oferecesse a possibilidade de melhorar a capacidade humana de lidar com tais problemas deveria incorporar a colaboração (GILLIES E CAILLIAU, 2000; BERNERS-LEE E FISCHETTI, 2000).

É válido notar, ainda, como enfatiza Berners-Lee e Fischetti (2000), que, ao contrário do Enquire, a ideia original da Web excedia a ideia de ser apenas um depósito de informações a serem extraídas, ou somente uma ferramenta de pesquisa e referência, ou, ainda, de comércio eletrônico. Estas são apenas partes da Web. “[...] *a generalidade era a essência de uma teia de informações*” (BERNERS-LEE e FISCHETTI, 2000, p.21).

2.2.2. A invenção: o processo

Em 1990, o projeto ganha um entusiasta dentro da CERN, um membro do *staff*, Robert Cailliau, que se ocupa de reescrever uma nova proposta em termos mais efetivos, intitulada “*WorldWideWeb: Proposal for a HyperText Project*”, realiza um lobby entre a vasta rede de amigos que possuía dentro da organização e tenta angariar estudantes ajudantes, dinheiro, máquinas e espaço físico para o desenvolvimento do projeto. Enquanto isso, Berners-Lee inicia a escrita do código da World Wide Web⁶⁶ de posse de um novo tipo de computador pessoal: o NeXT, produzido pela NeXT Inc⁶⁷, que trazia várias inovações como o primeiro “*point-and-click*” intuitivo, interface de pastas, entre outros. Assim, a justificativa para o desenvolvimento

65 Doug Engelbart é também o idealizador do *mouse*, que incorporava a ideia de que a interação entre indivíduo e máquina deveria ocorrer da forma mais natural possível. Segundo Berners-Lee e Fischetti (2000), “*Doug was too far ahead of his time*” (p.6). A revolução dos computadores pessoais que popularizaria o *mouse* viria quinze anos mais tarde. Ademais, como salientam Gillies e Cailliau (2000), o *mouse*, ou a tecnologia de “*point and click*” foi fundamental para a difusão da Web, contribuindo para o uso fácil e intuitivo. Suas ideias revolucionaram a forma de se pensar em documentos, a conexão entre eles e a interação entre humanos e computadores.

66 Este nome resultou de um processo de escolha guiado pela necessidade de que as palavras representassem o novo tipo de estrutura que se propunha: um sistema global de hipertexto. O nome refletia a natureza da distribuição dos indivíduos e computadores conectados pelo sistema.

67 Empresa fundada por Steve Jobs ao sair da Apple. Os computadores idealizados por ele eram destinados aos programadores, mas visavam oferecer a mesma facilidade de interatividade que a interface do Macintosh, além de incorporar tecnologias avançadas presentes nos computadores Alto da Xerox (GILLIES E CAILLIAU, 2000).

da programação do hipertexto global, sem o apoio formal da CERN, era experimentar o sistema operacional e o ambiente de desenvolvimento do NeXT⁶⁸.

Já existiam no mercado numerosos editores de hipertexto disponíveis, assim, o primeiro plano era usar um produto da indústria de hipertexto já disponível e apenas adicionar algum código de Internet para que os documentos de hipertexto pudessem ser enviados através da Internet. Obteve sucessivas respostas negativas que acompanhavam razões como ser muito complicado, caro, dificuldade de imaginar o projeto como um todo sem algo em funcionamento para demonstração e ausência de um nó central ao qual todos os demais estivessem conectados – pois eram assim que as coisas eram feitas. Então, Berners-Lee decide que precisaria escrever todo o código da Web e inicia em outubro de 1990.

O primeiro objetivo era escrever o que denominou por *Web cliente*, que era o programa que permitiria a criação, navegação e edição das páginas de hipertexto. O que se aproximaria, em aspecto, de um processador de texto, basicamente. O segundo passo era transformar o texto em hipertexto, o que era feito através da atribuição de um endereço de link de hipertexto para cada extensão do texto (que era um link). Posteriormente, escreve o código para o Hypertext Transfer Protocol (HTTP), que era a linguagem usada pelos computadores para comunicar através da Internet, e para o Universal Resource Identifier (URI), que era o esquema de endereçamento de documentos. Em dezembro de 1990, o primeiro *Web client* – um navegador/editor denominado WWW – já estava funcionando com o Hypertext Markup Language (HTML), escrito para descrever como formatar páginas contendo links de hipertexto. As páginas de Web, por sua vez, são mantidas em *Web server*, software que armazena as *Web pages* em uma seção do computador permitindo o acesso de outros indivíduos. O primeiro *Web server* escrito por Berners-Lee denominou-se “*info.cern.ch*”. Berners-Lee, então, cria a primeiro *Web page*, dentro deste servidor, contendo suas notas pessoais, especificações da HTTP, URI e HTML, e todas as informações relacionadas ao projeto⁶⁹.

68 “The NeXT interface was beautiful, smooth, and consistent. It had great flexibility, and other features that would not be seen on PCs till later, such as voice e-mail, and a built-in synthesizer. It also had software to create a hypertext program. Its failure to take over the industry, despite all these advantages, became for me a cautionary tale. NeXT required users to accept all these innovations at once – too much”. (BERNERS-LEE E FISCHETTI, 2000, p.28).

69 Para uma visão da primeira página de Web criada por Berners-Lee: <https://home.cern/science/computing/birth-web>.

A Web estava pronta para ser mostrada ao mundo, pelo menos em uma primeira aproximação do que viria a ser e do que foi idealizada. O problema é que ela somente funcionava na incomum plataforma do NeXT⁷⁰. O que criava um desafio adicional de incentivo ao uso generalizado: o grande incentivo para se disponibilizar uma informação na Web era a possibilidade de esta ser encontrada por qualquer indivíduo ao redor do mundo, mas como incentivar as pessoas a instalarem o programa se não havia informações interessantes o suficiente disponíveis na Web? E se apenas indivíduos dispostos de um NeXT podiam acessar as informações? O que urgia a criação de um browser cuja interface funcionasse em qualquer tipo de computador. A primeira versão desse navegador denominava-se *line-mode browser*, dado que foi pensado para ser tão simples que funcionasse em máquinas que só mostravam textos da forma de uma linha por vez. O problema com esse *line-mode client* era o fato de ele estar ainda em um estágio de desenvolvimento incipiente, permitindo apenas a navegação pela maioria das pessoas e não a edição de documentos.

Para resolver o problema de ausência de informações na Web, em quantidade e qualidade, Berners-Lee programa o browser para ser capaz de seguir links não somente a arquivos nos servidores HTTP, mas também em arquivos transmitidos pelo protocolo de Internet denominado FTP (*file transfer protocol*). Esses arquivos continham novos artigos e *newsgroup*. “Com essa mudança, artigos e grupos de notícias da Internet tornaram-se repentinamente disponíveis como páginas de hipertexto. De uma só vez, uma grande quantidade de informações que já estavam na Internet estava disponível na Web” (BERNERS-LEE E FISCHETTI, 2000, p.30)⁷¹.

O que havíamos conquistado até agora se baseava em alguns princípios-chave aprendidos por meio de experiências difíceis. A ideia da universidade era fundamental: a revelação básica era que um espaço de informação poderia incluir todas elas, dando enorme poder e consistência. Muitas das decisões técnicas surgiram a partir disso. A necessidade de codificar o nome ou endereço de cada objeto de informação em um URI era aparente. A necessidade de tornar todos os documentos de alguma forma “iguais” também era essencial. O sistema não deve restringir o usuário; uma pessoa deve ser capaz de se conectar com a mesma facilidade a qualquer documento, onde quer que esteja armazenado (BERNERS-LEE E FISCHETTI, 2000, p.33)⁷².

70 Gillies e Cailliau (2000) apontam que esse protótipo da WWW escrito no NeXT estabeleceu o padrão para os desenvolvimentos seguintes dado o nível de avanço em que estava. Ou seja, os demais projetos tentaram incluir todas as características e funcionalidades presentes na primeira versão, para os demais computadores. Como a capacidade de leitura e edição e uma visualização mais intuitiva e confortável com a tecnologia point-and-click.

71 Tradução do original em inglês.

72 Tradução do original em inglês.

Vale salientar que ao longo do processo acima descrito, o projeto aconteceu de forma informal dentro da CERN, nunca houve uma resposta das instâncias superiores, aos membros do *staff* aos quais a ideia era apresentada, as reações variavam de incredulidade à falta de entendimento do projeto como um todo. Alguns poucos estudantes, em programas provisórios de estágio, se engajaram no projeto durante a programação, mas tudo foi feito sem suporte formal interno, como fornecimento de pessoal, financiamento e estrutura física. Mesmo de posse do primeiro programa, funcionando plenamente entre alguns computadores da instituição, a resistência ao engajamento era enorme. Tal resistência era atribuída, por Berners-Lee, a necessidade de uma mudança de perspectiva quanto a interação com o computador.

Resumindo, a existência da Web marcava o fim de uma era de frustração onde os computadores não podiam se comunicar entre si devido suas incompatibilidades advindas da diversidade de fabricantes, de formatos de discos e de dados, esquemas de codificação de caracteres e redes incompatíveis. O que era imperativo para que a Web tivesse escala global, era definir poucas regras de protocolo que fossem básicas e comuns – regras para governar links de hipertexto entre servidores e *browsers* – que permitissem aos computadores a comunicação. Tais regras se circunscreviam, no caso da Web, às URIs, o esquema de endereçamento de cada informação a ser compartilhada; ao HTTP, o protocolo através do qual os computadores se comunicavam; e ao HTML, o formato para criação de documento de hipertexto que era exibido no *browser*.

Dada a ausência de adesão interna ao uso da Web, principalmente no que tangia suporte financeiro e de pessoal para o desenvolvimento de outros *browsers* que rodassem nos mais variados tipos de computadores e sistemas operacionais, Berners-Lee decide, em 1991, colocá-la disponível na Internet, seguida de uma postagem de aviso em todos os grupos de notícias possíveis na Internet. Assim, disponibilizando seu projeto para toda a comunidade acadêmica, não somente a comunidade da CERN, adeptos e entusiastas começam a surgir de várias regiões do mundo, criando servidores e Web *pages* fora da Europa, e, principalmente, interagindo via feedbacks que contribuíam para a sofisticação e maior desenvolvimento do programa. “O povo da Internet construiu a Web, de maneira verdadeiramente popular” (BERNERS-LEE E FISCHETTI, 2000, p.47)⁷³. Vale ressaltar que o que foi disponibilizado no servidor da CERN, criado por Berners-Lee, para livre download pela comunidade da Internet, foi um pacote de

73 Tradução do original em inglês.

softwares, chamado libwww e que consistia no protocolo da WWW dividido em subcomponentes, que permitia aos interessados escrever o código de seus próprios *browsers*.

2.2.3. A difusão: o início

Ficava claro que umas das razões para a lenta difusão do uso da Web era a ausência de outros browsers (ou Web *cliente*), tão intuitivos como o *point-and-click* desenvolvido para o NeXT, que funcionasse em outros computadores e sistemas operacionais, como PC, Macintosh, Unix. Dada a ausência de recursos para o projeto, internamente à CERN, Berners-Lee e Cailliau, iniciaram uma tentativa de estimular estudantes universitários da área de softwares a desenvolverem seus próprios protótipos de browser como um projeto integrante das atividades dos cursos. Estudantes da Helsinki University of Technology, escrevem um modelo para rodar no Unix operado pelo sistema X-Windows, denominado *Erwise*. Apesar de muito avançado, ainda não se enquadrava na categoria de editor. O Web browser ViolaWWW foi escrito por Pei Wei, na U.C. Berkeley e possuía muitos avanços: exibia HTML com gráficos, fazia animações e downloads de pequenos aplicativos da Internet. No entanto, apesar desses avanços, seu desenvolvimento não continuou no caminho da inclusão de funções de escrita e edição. No entanto, os diversos browsers começavam a se difundir. Nicola Pellow escreve Samba para Macintosh.

Sem um editor de hipertexto, as pessoas não teriam as ferramentas para realmente usar a Web como um meio íntimo de colaboração. Os navegadores permitiriam que eles encontrassem e compartilhassem informações, mas eles não poderiam trabalhar juntos de forma intuitiva. Parte do motivo, imaginei, era que a colaboração exigia muito mais uma mudança social na maneira como as pessoas trabalhavam. E parte disso era que os editores eram mais difíceis de escrever (BERNERS-LEE e FISCHETTI, 2000, p.57)⁷⁴.

Na medida em que mais opções de *browsers* ficavam disponíveis, novos servidores também, com uma frequência cada vez maior. E servidores voltados para novas formas de se fazer as coisas apareciam e reforçavam o entusiasmo com o campo que nascia. Um exemplo foi o primeiro museu virtual criado por Frans van Hoesl, cujo trabalho se constitui em criar um Web site de hipertexto para um material disponível em formato de arquivos de imagens em um servido de Internet com protocolo FTP. Esse material era uma compilação de fotos de uma exibição sobre Roma durante o Renascentismo organizada pelo Vaticano no Congresso Norte

74 Tradução do original em inglês.

Americano. Inovador também foi o servidor criado por Steve Putz para mapas, que derivavam de uma base de dados de informações geográficas. “Esse uso da Web para levar pessoas distantes a grandes recursos e o idioma de navegação usado para fazer o museu virtual se popularizou e inspirou muitos sites excelentes” (BERNERS-LEE E FISCHETTI, 2000, p.59)⁷⁵.

Com o surgimento de uma crescente competição entre os *browsers*, começaram, então, a surgir desenvolvedores comerciais e institucionais em cena. Como o *browser* Arena, criado por Dave Raggett da Hewlett-Packard; o Lynx 2.0 desenvolvido por Lou Montulli da *University of Kansas*; e o Mosaic, escrito por Marc Andressen e Eric Bina, ambos membros do grupo de design de software do *National Center for Supercomputing Applications (NCSA)* da *University of Illinois at Urbana-Champaign*. Este último, lançado na Web em 1993, obteve uma disseminação mais rápida do que os demais browsers, dada suas características que tornavam o uso mais fácil e intuitivo, principalmente por ser o resultado de um processo de desenvolvimento intimamente correlacionado com os feedbacks dos usuários de browsers. Gillies e Cailliau (2000) apontam que o lançamento do Mosaic em 1993 é, em grande medida, responsável pela definitiva difusão da Web fora do meio acadêmico, pois rodava na maioria dos computadores em uso pela grande parcela da população (Macintosh, PCs, etc). Ademais era fácil de instalar, posteriormente permitiu a inclusão de imagens *inline* e os bugs relatados pelos usuários eram rapidamente corrigidos. Fácil acesso para iniciantes na Web.

No entanto, percebeu-se uma resistência do mercado à Web dado o temor da existência de licenças de qualquer tipo, configurando restrições ligadas ao custo. Assim, segue-se a decisão do inventor e da CERN de colocar a tecnologia da Web em domínio público geral, o que permitia que qualquer pessoa usasse o protocolo e código da Web livre de encargos, criar um servidor ou um browser, doa-los ou vende-los, sem royalties ou outras restrições.

Assim, surge a demanda pela criação de uma organização que monitorasse o desenvolvimento da Web: “A Web estava começando a mudar de fase”, (BERNERS-LEE e FISCHETTI, 2000, p.75)⁷⁶. Com o percebido crescimento exponencial de acessos ao servidor da CERN e de novos servidores e o aumento do tráfego na Web, era importante dar segurança às companhias, principalmente às ligadas ao setor de computação, de que a Web iria evoluir de forma estável.

75 Tradução do original em inglês.

76 Tradução do original em inglês.

Eu não queria formar um corpo de padrões per se, mas algum tipo de organização que pudesse ajudar os desenvolvedores ou servidores e navegadores a chegar a um consenso sobre como a Web deveria operar. [...] a evidência dizia que “a Web” poderia se dividir em várias facções - algumas comerciais, algumas acadêmicas; alguns gratuitos, outros não. Isso iria contra o próprio propósito da Web: ser um meio de hipertexto único, universal e acessível para compartilhar informações (BERNERS-LEE e FISCHETTI, 2000, p. 76).⁷⁷

Ademais, a *European Commission (Fraunhofer Gesellschaft)*, juntamente com a CERN, inicia o primeiro projeto, da União Europeia, baseado na Web, denominado Webcore, para disseminação de informação tecnológica para países europeus do antigo bloco Soviético. Assim a mídia, em dezembro de 1993, toma conhecimento e começa a publicar massivamente sobre a Web e o Mosaic, potencializando a disseminação de informações sobre as novas tecnologias.

Em 1994 começam a surgir companhias ofertantes de serviços de Internet, como a empresa O’Reilly que criou um produto denominado “*Internet in a Box*” que objetivava trazer a Internet e a Web para os lares:

Já era possível para qualquer pessoa baixar, gratuitamente, todos os navegadores, TCP / IP e softwares necessários para entrar na Internet e na Web, mas o usuário precisava saber muito sobre como configurá-los e fazer com que funcionassem juntos, o que foi complicado. Nem a Internet nem a Web foram inicialmente configuradas para uso doméstico ou empresarial individual; eles foram feitos para universidades, pesquisadores e organizações maiores. O produto da O’Reilly juntou tudo. Tudo o que o usuário precisava fazer era instalá-lo em seu computador e pagar as tarifas telefônicas de sua conexão com a Internet (BERNERS-LEE e FISCHETTI, 2000, p.80)⁷⁸.

Em abril de 1994 tem-se a criação da Netscape com o lançamento de seu novo browser, Mozilla, em outubro do mesmo ano. A Netscape simboliza o surgimento de softwares e serviços da Web como produtos comerciais, e é um passo importante para a Web. A Web torna-se um negócio. Na mesma época, a Microsoft decide que a próxima versão de seu sistema operacional, Windows 95, incluiria um software de acesso à Internet e que esta constituiria uma nova e importante parte da estratégia da companhia. “Em vez de desenvolver seu próprio código, a Microsoft licenciou o código do navegador de uma pequena spin-off do NCSA chamado

77 Tradução do original em inglês.

78 Tradução do original em inglês.

Spyglass. O custo foi de US \$ 2 milhões - mais dinheiro do que qualquer um de nós envolvidos desde os primeiros dias jamais teria sonhado” (BERNERS-LEE E FISCHETTI, 2000, p.93)⁷⁹.

Como resposta à demanda de criação de uma organização que monitorasse o desenvolvimento da Web, surge o *World Wide Web Consortium* (W3C), instaurado no *Laboratory for Computer Science* do MIT. Faziam parte do W3C os diversos membros de organizações como Netscape, Microsoft e IBM. Os membros poderiam ser de organizações comerciais, educacionais ou governamentais. O propósito do consortium seria “levar a WEB a todo o seu potencial” (BERNERS-LEE E FISCHETTI, 2000, p.94)⁸⁰, desenvolvendo, primeiramente, protocolos comuns para melhorar a interoperabilidade e a evolução da Web. “Para fazer isso, ficaríamos à frente de uma onda significativa de aplicativos, serviços e mudanças sociais, cumprindo uma combinação única de funções tradicionalmente atribuídas a organizações bastante diferentes” (BERNERS-LEE E FISCHETTI, 2000, p.94)⁸¹. O W3C desenvolveria especificações técnicas, ajudaria a desenhar e desenvolver códigos, representaria o poder e autoridade de milhões de desenvolvedores, pesquisadores e usuários e alavancaria os avanços mais recentes na tecnologia da informação. O consortium seria financiado pelos pagamentos de membros, mas também recebia financiamento dos governos norte-americanos e europeus. Por exemplo, a agência de defesa dos EUA, por meio de seu setor de projeto de pesquisas avançadas, financiava o consortium por este promover uma ponte entre pesquisa acadêmica e indústria.

Uma das primeiras tarefas do consortium (a lista de prioridade era criada a partir dos feedbacks dos membros) seria lidar com a questão da segurança da rede, ou seja, como proteger as informações dos usuários, como dados de cartão de crédito. Resolver essa questão seria de especial importância para permitir o e-commerce – o comércio eletrônico. Antecipando-se ao W3C, dada a percepção da importância para os futuros desenvolvimentos da empresa e de um nicho de mercado, a Netscape lança um software, o *Secure Sockets Layer*, destinado a proteger os dados pessoais dos usuários na Web.

Em dezembro de 1994 três eventos tomam lugar e alteram definitivamente o futuro da Web a saber: o primeiro encontro dos membros do W3C; o lançamento, pela Netscape, da versão

79 Tradução do original em inglês.

80 Tradução do original em inglês.

81 Tradução do original em inglês.

comercial do seu browser. O Mozilla (Navigator 1.0) é lançado livre na Internet, sem cobranças, dada a percepção da Netscape da urgência de lançar o browser antes das demais companhias, como a Microsoft. O modelo de negócio – a forma de lucrar – seria por meio de publicidade exibida no site da Netscape - quando os indivíduos se dirigissem ao site para fazer o download do browser ou quando abrissem o browser, pois o site da empresa seria a página padrão ou de inicialização -, pela disponibilidade de outros serviços para venda no site, e pelo fornecimento de navegadores comerciais para companhias, assim como a criação e hospedagem de servidores, entre outros). “A Netscape estava sabiamente reconhecendo isso na Web. Era mais lucrativo ser uma empresa de serviços do que uma empresa de software” (BERNERS-LEE e FISCHETTI, 2000, p.100)⁸². O modelo de negócios da Netscape definiu o caminho que as demais companhias da Web seguiram: lançar uma versão básica do software livre na Internet para atrair o máximo de usuários possíveis – que eventualmente fariam sugestões e feedbacks para melhora do software – e fazer dinheiro através das milhares de visitas ao site através de anúncios e serviços.

O terceiro evento foi a decisão, por parte da CERN, de não ser o escritório, ou a base, do consortium na Europa. Esse lugar foi, então, assumido pelo *Institut National Recherche en Informatique et en Automatique (INRIA)*, localizado na França.

Um sinal da difusão global e importância da Web era o reconhecimento, pelos governos ao redor do mundo, da influência da nova tecnologia a ponto de virar um tópico de discussão no encontro anual do G7 em fevereiro de 1995 em Bruxelas. Nesse meio tempo, a Microsoft tenta licenciar o browser da Netscape e incorporá-lo ao seu sistema operacional Windows 95, mas é recusada e redobra seus esforços na busca de seu próprio browser. Em abril de 1995, a Compaq anuncia que sua nova linha de computadores pessoais viria com o Navigator da Netscape, configurando a primeira vez que um browser viria junto com o hardware. Em maio, Sun Microsystems lançam o Java, uma nova linguagem de programação⁸³, o que abre um mundo de possibilidades de aplicações para a Web. Netscape imediatamente licencia o Java e incorpora em seu browser, Navigator. A vantagem do Java era possibilitar computadores com discos

82 Tradução do original em inglês.

83 “*Small application programs written in Java, called applets, could be sent directly between computers over the Internet, and run directly inside a Web page on a browser. [...] It met the need for applications in which a hypertext page was not sufficiently interactive, and some programming on the client was necessary.* (p.104)

rígidos e RAM com menores espaços, memória, o que criava a possibilidade de computadores pessoais mais baratos.

Outro evento sintomático da importância e difusão da Web foi abertura de capital da Netscape, dada a necessidade de uma grande infusão de capital que permitisse a recém-lançada empresa, apenas 16 meses de existência, competir com gigantes como a Microsoft. A “*initial public offering*” (IPO) da Netscape foi considerada a maior da história até então, principalmente quando se considera que a empresa ainda precisava demonstrar algum lucro. Após um único dia de negociação na bolsa de valores, Netscape valia \$4.4 bilhões. O que retratava as elevadas expectativas de Wall Street com relação às ações de alta tecnologia em 9 de agosto de 1996.

Se a World Wide Web ainda não tinha chamado a atenção total do público, esta história notável a colocou no centro do palco. Também enviou uma mensagem inegável ao mundo comercial: a Web era um grande negócio. A corrida do ouro começou. A enxurrada de dinheiro permitiu que a Netscape comprasse pequenas empresas que desenvolveram produtos especializados para a Web, criasse joint ventures com corporações maiores e ampliasse sua linha de produtos para dar suporte a grandes contratos de grandes compradores corporativos. No final de 1996, quando se estabeleceu em seu modelo de negócios completo, a Netscape empregaria mais de duas mil pessoas e relataria receitas de US \$ 346 milhões. O preço inflacionado de suas ações cairia para níveis razoáveis nos próximos anos, mas de uma só vez a Web se tornou um grande mercado (BERNERS-LEE e FISCHETTI, 2000, p.106)⁸⁴.

Duas semanas após a IPO da Netscape, Microsoft lançou o Windows 95 com seu próprio browser, o Internet Explorer. Deixando claro a definição da companhia na direção da Internet.

Segundo Berners-Lee e Fischetti (2000) e Gillies e Cailliau (2000), o que se segue, então, é o início de uma alteração da dinâmica da economia, das instituições e da sociedade em torno das novas tecnologias de informação e conectividade. Surgem novas empresas de tecnologia da informação e novas mercadorias, empresas já estabelecidas percebem a necessidade de reorientar suas trajetórias – caso da Microsoft e da IBM -, as instituições buscam formas de acompanhar os avanços inovativos etc. Leis e processos judiciais marcam as primeiras décadas do surgimento da Web no que tange o controle da informação disponível e o acesso a ela. E debates sobre regulações de mercado para garantir a competitividade são alguns exemplos das perturbações causadas pelo que parece ser uma revolução nos modos de vida. “As lutas competitivas por pedaços de um mercado lucrativo agora forneciam o pano de fundo financeiro

84 Tradução do original em inglês.

para a revolução tecnológica, que por si só era o pano de fundo para uma verdadeira revolução social” (BERNERS-LEE e FISCHETTI, 2000, p.112)⁸⁵.

2.2.4. O surgimento de novas indústrias, setores de atividade e modelos de negócios: o caso da Google.

A invenção e difusão da WWW possibilitou o surgimento de inúmeras novas tecnologias e o aumento das possibilidades de seu uso por indivíduos não especialistas. Após sua consolidação, observa-se um crescimento do número de websites na rede, criando uma demanda por mecanismos que permitissem encontrar as informações específicas procuradas pelos usuários individuais e corporativos, em parte devido ao caráter descentralizado da Web (GREENSTEIN, 2015; ALBUQUERQUE, 2019). Nesse contexto surge, em 1998, a Google, uma empresa criada a partir de um algoritmo para buscas na Internet, esta, por sua vez, formalizada como uma das novas mercadorias dessa nova era. A Google foi fundada por Sergey Brin e Lawrence Page, estudantes do programa de Phd da *Stanford University* em ciências da computação. Sua invenção, que levou à fundação da Google, se resumia em um mecanismo de busca de larga escala. Ambos eram participantes de um projeto de pesquisa financiado pela *National Science Foundation* (NSF) denominado “*Digital Library Initiative*” cujo um dos objetivos era melhorar a ciência de recuperação e armazenamento de informações em grande escala. Assim, o objetivo de pesquisa de Brin e Page, a princípio, era inventar ferramentas para examinar o conteúdo digital de bibliotecas. O início do projeto se dá em 1994, quando a Web ainda está em seus primeiros momentos de difusão e, com o crescimento acelerado de seu uso ao longo do globo, o interesse de Page se volta para o conteúdo disponível na Web e propõe examinar as informações contida nas *web pages*. Assim, o foco da pesquisa se altera, direcionando-se para a invenção de ferramentas que examinassem o conteúdo da Web.

Dada a potencial novidade científica nessa proposta, os coordenadores do projeto permitiram que essa mudança de caminho na pesquisa fosse efetuada. Assim, Brin e Page desenvolveram um método para mensurar os links entre websites, gerando uma classificação destes segundo a frequência de links que cada website recebia, ou seja, quanto maior a quantidade de links um website recebesse, maior seria sua posição na classificação. Segundo Greenstein (2015), o algoritmo criado mensurava “a popularidade de uma página da Web” (p.368). A vantagem

85 Tradução do original em inglês.

dessa classificação era permitir melhores respostas às consultas sobre como encontrar *web pages* informativas, dado que estas, em geral, recebiam mais links e, portanto, à esta tecnologia deu-se o nome de mecanismo de busca. O algoritmo criado por Page e Brin recebeu o nome de Page-Rank, que foi a primeira versão do que viria a ser o Google.

Percebe-se aqui o papel da Web em direcionar a pesquisa científica, dando a ela outros rumos além dos inicialmente propostos. Ou seja, essa mudança de rumo de pesquisa só foi possível pela invenção e posterior difusão da Web, dada a criação de uma necessidade: encontrar informações na avalanche de *web sites* que crescia.

Em seguida, implementaram o Page-Rank em um mecanismo de busca, denominado Backrub, disponibilizado em um servidor da Stanford, o que permitiu o refinamento do algoritmo através de feedbacks de usuários. Dada a popularidade do mecanismo de busca entre os usuários dentro do campus, os inventores partiram para uma tentativa mercadológica: receber alguma renda através de licenciamentos para outras entidades e firmas. Nesse interim, a tecnologia é patenteada tendo a Stanford como titular da patente e o inventor como Lawrence Page. A princípio, encontram dificuldades em licenciar o algoritmo, não encontrando muito interesse, pois era uma época de efervescência de *startups* que objetivavam lucrar com o crescimento da Web e existiam inúmeras outras abordagens de busca na Internet disponíveis, e sua proposta, a princípio, não parecia se distinguir das demais. Outros mecanismos de busca com intuição similar eram o RankDex, criado por Robin Li que fundou a Baidu, e o Inktomi, de Eric Brewer e Paul Gauthier, considerado o que dominava o mercado no final da década de 1990.

Um dos princípios fundamentais da Web, como mencionado anteriormente, era a descentralização ou ausência de um controle central: qualquer usuário da Internet poderia adicionar informações e criar links entre estas. Essa característica proporcionou uma oportunidade de negócios que se baseava em ajudar os usuários a encontrar o que desejavam no crescente conjunto de *web pages*, ou seja, uma ajuda para navegação. Uma das primeiras tentativas de explorar esse nicho de mercado se traduziu na tentativa de cobrar por assinaturas de informações organizadas da Web, uma ideia parecida ao catálogo telefônico, porém, não obteve sucesso. Outra abordagem se materializou na existência dos grandes portais, como Yahoo e AOL, que se baseavam na noção de interesses coletivos compartilhados, voltados para grupos de usuários, oferecendo ampla gama de conteúdo específico e selecionado. Esses serviços eram prestados por intermédio das rendas advindas, em boa parte, da publicidade

exibida nesses portais que, por sua vez, prometiam entregar, para as empresas anunciantes, uma mensuração dos resultados dos anúncios e melhor segmentação do público – o que era essencial para o anunciante e para o setor de publicidade como um todo.

Uma terceira abordagem se concentrava em oferecer uma solução técnica ao problema de como navegar nas milhares de páginas disponíveis na Web ou como encontrar o que se desejava de forma eficiente. Assim, começam a surgir soluções de engenharia voltadas para a pesquisa online. A oferta desse serviço, em geral, estava atrelada à venda de resultados das buscas a empresas que licenciavam o serviço ou à venda de publicidade, o que se mostrou mais atraente e com maior potencial lucrativo. E é nessa área que a Google surge e atua. Apesar do Cookie⁸⁶ – uma tecnologia de rastreamento da atividade do usuário na Web ou de criação de histórico do navegador - já estar disponível desde 1994 e ter sido implantado por muitas firmas, como os grandes portais, criando a possibilidade de melhor desenho das preferências dos usuários e, portanto, melhor segmentação, quando a Google surge, as firmas ainda encontravam considerável dificuldade em anunciar com precisão. Além do mais, a prática comum de negociação dos anúncios, os banners, se baseava no preço de impressão não estando diretamente relacionado ao seu impacto, o que evidenciava o problema de mensuração.

Assim, a Google surge em um contexto de elevada oferta de mecanismo de buscas e reduzida demanda, em um mercado dominado pelos portais que não viam muita utilidade nestas tecnologias e, adicionalmente, consideravam-nos inconveniente por, eventualmente, direcionar o usuário para fora do portal, o que diminuía a possibilidade de venda de anúncios que era baseada no volume de tráfego de navegação. No entanto, segundo Greenstein (2015) a empresa possuía alguns diferenciais que, inicialmente, ajudaram a atrair a atenção do mercado e dos usuários em geral, como um design simples, que destoava do visual dos portais, e uma alegada neutralidade comercial com respeito a publicidade. O foco do serviço oferecido era entregar resultados relevantes e rápidos. Inicialmente ofereceu apenas um serviço de busca que se resumia a uma barra para inserção de uma palavra-chave pelos usuários, que, por sua vez, gerava um conjunto de resultados que a empresa devolvia ao usuário. Começaram por

⁸⁶ Essa tecnologia permitia que informações do navegador do usuário fossem acessadas pelo website criando um histórico de navegação, o que criava uma base de dados com valor inestimável. Seu uso beneficiou grandes firmas com massiva quantidade de conteúdo e aquelas com capacidade para rastrear os usuários entre múltiplos sites. Dando também origem ao rastreamento por terceiros, ou seja, extrapolava a esfera website rastreando seu usuário. Um ramo de rastreamento universal de usuários. Criando um comércio do surfing.

disponibilizar o serviço – o mecanismo de busca - a qualquer usuário, o que serviu ao objetivo de elevar a parcela do mercado e conseguiram captar receitas advindas da exibição de anúncios simples e banners, que eram exibidos por um período fixo de tempo e que não adaptavam o conteúdo ao usuário. Mantendo assim seu alegado compromisso com a neutralidade comercial: criando uma separação entre os anúncios e os resultados de pesquisa e rejeitando o “*pay-to-list*”.

Com o crescimento do número de usuários do mecanismo de busca da Google, os portais começaram a rever seus contratos com as demais empresas de busca na Internet. A Google fecha seu primeiro contrato com um cliente pagante em junho de 2000, a Yahoo, o que representava uma elevação de receitas, da legitimação no mercado e do número de usuários de forma considerável, permitindo, inclusive e primordialmente, que o algoritmo fosse refinado. A importância desse acordo com a Yahoo pode ser corroborada, também, pelos novos 130 acordos comerciais, com outras firmas, feitos nos dois anos posteriores. Inicialmente, os portais não enxergavam empresas como a Google, fornecedoras de mecanismos de buscas, como concorrentes diretas, dado que estes ofereciam linhas adicionais de negócios, como e-mails, grupos de discussão e notícias. Assim, os mecanismos de busca eram vistos mais como um complemento em seus sites. Como “parecia não representar uma ameaça competitiva imediata como substituto”, (GREENSTEIN 2015, p.378), houve cooperação e envio de tráfego por parte dos grandes portais para o serviço da Google.

A Google se preocupava em fazer investimentos massivos e continuados para melhoras técnicas, como, por exemplo, em tecnologias de hardware de servidores para alcançar e manter a rapidez de processamento, uma de suas marcas, e aos poucos evoluiu para uma plataforma de publicidade online. Em 2001 adotou a abordagem de leilão de palavras chaves ligadas ao mecanismo de busca – que já era utilizado por outras empresas -, mas não alterou a lista orgânica, gerada pelo algoritmo, para atender aos interesses dos anunciantes, identificando os anúncios ao listá-los acima e à direita dos resultados da pesquisa e indicando-os como pagos. Inovou ao desenhar um mecanismo de leilão de segundo preço, para evitar a manipulação do leilão pelos anunciantes, e a exibição dos anúncios obedeciam às seguintes regras: os anunciantes pagariam por click dos usuários nos anúncios exibidos e não pela “impressão”, o que contribuía para a melhora da mensuração do alcance da publicidade; os anunciantes não precisariam pagar se o anúncio não atraísse clicks dos usuários.

Assim, a Google desenvolve o “*quality ranking*” baseado no engajamento dos usuários aos anúncios e resultados de pesquisa. Esse mecanismo era utilizado nos leilões, o que funcionava como um filtro, apurando a segmentação do conteúdo a ser anunciado. Criava incentivos para melhoria constante dos websites, uma vez que as empresas buscariam melhorar o score tanto para o leilão quanto para melhorar sua classificação na lista orgânica do mecanismo de busca. O que aprimorava os resultados entregue aos usuários, aumentando o número de tráfego. Ou seja, um ciclo virtuoso de *feedback*. Isso só foi possível, mais uma vez, devido à investimentos massivos e constantes em ciências da computação, segundo Greenstein (2015), a empresa precisou participar da invenção de uma ciência da computação extremamente complicada para implementar o “*quality ranking*”: “a ciência da computação de um algoritmo de classificação de qualidade automatizado para toda a Web é extremamente desafiadora do ponto de vista técnico e impossível para muitas organizações” (GREENSTEIN, 2015, p.383-4).

Em 2003, já considerada liderança na publicidade, passa a oferecer um novo serviço de exibição de anúncios em lugares específicos dentro de blogs – em geral, pequenas janelas retangulares, o que foi denominado “*Ad Sense*”. Este era fornecido através de um leilão com mecanismo *Vickery-Calrk-Groves*, no qual o lance era lacrado e se cobrava do licitante o dano que, porventura, causasse aos demais participantes. Para dispor os anúncios nas diversas páginas da Web, a Google dividia as receitas dos anúncios com os sites que os exibiam. Em resumo, a grande inovação da Google foi o “*quality-weighted second-price keyword position auction for ads*”, que fornece ao mercado de publicidade a mensuração e segmentação eficaz e aos usuários, resultados relevantes de suas buscas na Internet. Sofisticando e dominando o mercado. Torna lucrativa a Web, atrai usuários, interesses e atenção para o potencial de negócios, lucros e mercados. Seu foco na melhoria contínua atrai cada vez mais tráfego: um exemplo é a tecnologia de *spell checking* que dobra o tráfego por facilitar a navegação. Ainda, segundo Greenstein (2015): “A ascensão do Google solidificou o movimento irreversível de afastamento das raízes inocentes da Web em direção a uma comercialização generalizada do surfe” (GREENSTEIN, 2015, p.386)⁸⁷.

Segundo Albuquerque (2019), a comercialização dessa mercadoria envolve a relação entre os usuários da WWW e os anunciantes, o que, por sua vez, determinou uma profunda transformação no mercado de propaganda, no qual, os gastos na Web ultrapassaram o total

87 Tradução do original em inglês.

investido nos meios tradicionais. Com as novas empresas capturando a maior parte desses gastos, sendo a Google a ocupante do primeiro lugar. Ademais, é relevante o surgimento de um conjunto de inovações relativas à forma como a Google precifica os seus anúncios, através da utilização de técnicas de leilão. A busca na Internet é custeada por gastos em propaganda, dessa forma, estabelece uma dinâmica específica onde a coleta de dados dos usuários, para determinar com maior precisão o público-alvo das propagandas e, assim, cobrar mais pelo serviço oferecido, é aprofundada.

2.3. Considerações finais.

De acordo com Albuquerque (2019), a existência de um conjunto de universidades com capacitação científica e tecnológica em diversas áreas, especialmente ciências da computação, de firmas com produtos e departamentos de P&D capazes de atender a uma demanda em construção sobre redes e sobre conectividade, de milhares de profissionais capazes de implementar o funcionamento da rede em construção e impulsionar o gigantesco processo com vários momentos de invenção coletiva, formaram o contexto de surgimento da WWW. Segundo Dertouzos (2000), muitos cientistas da computação se debruçaram, ao longo das décadas anteriores a invenção da WWW, sobre dois tópicos, hipertexto e redes de computadores, mas somente Berners-Lee concebeu uma forma de juntar os dois elementos para criar a Web.

Ademais, a invenção da WWW dependeu de um amplo conjunto de inovações anteriores na indústria de computadores, a saber: novos produtos e processos gerados a partir da emergência da computação eletrônica, em meados do século XX. Com a difusão do uso de computadores, surge uma demanda sobre como conectar diferentes computadores e seus usuários. A invenção da WWW, então, permitiu a utilização da Internet de forma muito mais intensa e mais ampla, uma pré-condição para a transformação comercial da internet. No entanto, o crescimento da rede em si exige novos tipos de computadores, servidores e roteadores. A expansão da WWW se associa com a expansão no uso de computadores pessoais. Atualmente, inovações como o desenvolvimento de smartphones abriu outras áreas de expansão da WWW.

Albuquerque (2019) afirma que a invenção da WWW teria características de uma inovação radical, localizada na origem de uma *big-bang* de uma nova revolução tecnológica, pelo enorme impacto causado em inúmeros setores da atividade econômica. E o que foi relatado acima

parece confirmar que o resultado da invenção da WWW e da cadeia de eventos – inovações complementares, efeitos para a frente e para trás nas cadeias industriais e econômicas – por ela desencadeada é o estabelecimento de uma nova estrutura na economia mundial.

CAPÍTULO 3: PROPOSTA METODOLÓGICA PARA ANÁLISE DA TRAJETÓRIA DA *WORLD WIDE WEB*.

Com o objetivo de analisar o comportamento da WWW, bem como do sistema complexo que esta inovação conformou, a fim de identificar uma dinâmica que se assemelhe à de uma GPT, as próximas subseções apresentarão uma proposta de investigação desta dinâmica através de metodologias empíricas. Para tanto, propõe-se a construção e análise de duas redes, uma de patentes e outra de artigos científicos, compostas por suas respectivas citações, para captar as dinâmicas de ciência e tecnologia deflagradas a partir da invenção da WWW. Esta proposta se baseia nos desenvolvimentos recentes das metodologias de análise de redes usadas para a representação de sistemas complexos. Desta forma, as duas primeiras seções compreendem uma revisão teórica e metodológica sobre o uso de patentes e artigos, bem como de suas redes de citações, como indicadores da dinâmica inovativa e científica. A terceira se dedica a uma apresentação sucinta de alguns trabalhos na literatura que se propuseram a identificar tecnologias GPTs através do uso de dados de patentes e artigos. Por sua vez, a quarta seção descreve as redes que serão utilizadas e a metodologia utilizada em suas construções. E a quinta e última seção se ocupará da apresentação do método a ser utilizado para análise das redes, a saber a ciência de redes.

3.1. Patentes e citações de patentes como indicadores: uma revisão conceitual e metodológica.

Patentes têm sido largamente consideradas como uma manifestação empírica do resultado da atividade inovativa, e, nas últimas décadas, a utilização de suas estatísticas em pesquisas sobre a dinâmica do progresso tecnológico, como potenciais indicadores do processo inovativo, vem se aprofundando (JAFFE E TRAJTENBERG, 2002). Assim, também, como na tentativa de capturar a importância do conhecimento tecnológico e seu *spillover*, bem como seus efeitos sobre a economia. A tecnologia ou o conhecimento tecnológico pode ser definido como um conhecimento aplicado à uma, ou em uma, técnica. Assim, as patentes são reconhecidas como uma forma de mensurar a produção de conhecimento tecnológico e sua dinâmica dado que o agente inovador ao visar o lucro, possui como incentivo para inovar o monopólio de uso dessa técnica (ou tecnologia) conferido pela patente. Por isso, é uma forma de avaliar o resultado do

processo inovativo, ou de mensurar o resultado dos esforços de P&D, ou, ainda, de captar a dinâmica inovativa pelo lado da oferta.

A patente é um instrumento de apropriação da produção intelectual que resulta do esforço inovativo em geral, advindo de atividades bem estabelecidas de P&D nas firmas e universidades. Por isso, muitas vezes, são tidas como indicadores do “lado da oferta” do processo inovativo. Garantem o direito de propriedade da invenção, descrita no documento de patente, ao seu titular. O documento de patente, em si, contém uma quantidade considerável de informações sobre a invenção reivindicada, tornando-o uma fonte valiosa de dados para trabalhos empíricos sobre o progresso tecnológico. Informações como autoria ou inventor (1º autor representa a principal localização da atividade tecnológica), titularidade (o detentor dos direitos exclusivos de produção, comercialização e licenciamento da invenção, podendo ser uma organização, por exemplo), endereço do inventor, país, classe tecnológica, citações à patentes anteriores e citações não patentes (citações à artigos e publicações científicas, entre outros) e uma descrição completa da invenção reivindicada pela patente compõe o escopo do documento de patente. A patente concedida é um documento público, emitido por um escritório governamental autorizado, que garante direitos de propriedade intelectual exclusivos por um limitado período de tempo (GRILICHES, 1990 e 1984; JAFFE E TRAJTENBERG, 2002; TRAJTENBERG, 2002; HALL, JAFFE E TRAJTENBERG, 2002; NAGAOKA, MOTOHASHI E GOTO, 2010).

Para que uma patente seja concedida, é preciso que a invenção reivindicada obedeça aos critérios de concessão dos diversos escritórios de patentes nos quais fora depositada, os padrões de novidade e potencial utilidade, que podem apresentar diferenças de abordagem de um escritório para outro. Dessa forma, a taxa de concessão e o rigor da avaliação, também variam entre países e ao longo do tempo. Tal variação é associada às diferenças nos procedimentos e recursos dos variados escritórios de patentes, podendo implicar em diferenças na qualidade média de uma patente concedida entre países e períodos (GRILICHES, 1990). Portanto, há uma diferença entre o número de depósitos de patentes e o de concessão. Sabe-se ainda, que o número de depósitos pode variar de acordo com as condições econômicas e, geralmente, se observa um decréscimo na taxa de depósito durante períodos de depressão econômica, como crises e guerras. O propósito do sistema de patentes seria encorajar a invenção e o progresso técnico através da promoção de um monopólio temporário e por forçar a divulgação precoce da

informação necessária para a produção do item ou a operação do novo processo (GRILICHES, 1990; HALL, JAFFE E TRAJTENBERG, 2002).

As citações feitas às patentes prévias, constantes no documento de patente, que têm especial relevância para este trabalho, e seu uso em trabalhos empíricos tem aumentado significativamente, dado que possuem um papel legal no documento: definir os limites da proteção que está sendo reivindicada pelo titular. Ou seja, identificar o estado da arte tecnológico sobre o qual tal invenção se debruça e qual sua real contribuição. As citações de patentes, no fim, são resultado de um trabalho feito pelo solicitante da patente e pelos examinadores dos escritórios de patente, e tal processo parece fornecer os incentivos corretos para que todas as patentes relevantes e necessárias sejam citadas e apenas elas (TRAJTENBERG, 2002). Sobre esse fato se fundamenta o uso das citações para inferir a importância tecnológica da invenção subjacente, bem como seu spillover. Segundo Jaffe e Trajtenberg (2002), usar uma base de dados constituída por citações de patentes cria a possibilidade de traçar múltiplos links entre invenções, inventores, cientistas, firmas, localização etc.

Em particular, citações de patentes permitem o estudo de spillovers, e a criação de indicadores da importância ou impacto tecnológico de patentes individuais, assim introduzindo uma forma de capturar a enorme heterogeneidade no valor das patentes (JAFFE E TRAJTENBERG, 2002)⁸⁸.

A concepção do papel das citações de patentes está baseada na visão cumulativa do processo de desenvolvimento tecnológico, pela qual cada inventor se beneficia do trabalho de inventores anteriores e, por sua vez, contribui para a base de conhecimento sobre a qual futuros inventores irão construir. A premissa é a de que uma citação de patente constitui um importante sinal da relação tecnológica entre as invenções citantes e as citadas. A partir dessa premissa, são feitas hipóteses sobre como o processo cumulativo de desenvolvimento tecnológico deve se manifestar nos dados de citação: “usamos o número e o caráter das citações finalmente recebidas por uma determinada patente para caracterizar o impacto tecnológico e econômico de uma determinada invenção” (JAFFE E TRAJTENBERG, 2002, p.55)⁸⁹.

88 Tradução do original em inglês.

89 Tradução do original em inglês.

No entanto, alguns problemas, bem como algumas especificidades, podem surgir do uso de dados de patentes e de suas citações como indicadores do processo inovativo: o primeiro tange a representatividade dessa forma de apropriação intelectual, ou seja, dada a existência de outras formas de proteção intelectual, que podem servir, mais ou menos, aos objetivos e estratégias dos diversos agente inovativos, provavelmente, o uso exclusivo de patentes como indicador da produção tecnológica pode subestimá-la. Ou ainda, alguns setores podem estar mal representados por esses indicadores. Além do mais, nem todas as invenções satisfazem aos critérios de patenteabilidade dos diversos escritórios⁹⁰. Dessa forma, é preciso ponderar e explorar sobre a extensão da representatividade da patente para o amplo universo de invenções (JAFFE E TRAJTENBERG, 2002; GRILICHES, 1990). O uso da primeira classe tecnológica que aparece no documento de patente também cria algumas dificuldades, dado que podem existir mais de uma classe e subclasse – e essas, em geral, variam entre os escritórios - e que a classificação não está em ordem de relevância, permitindo acompanhar apenas os macro movimentos.

Griliches (1990) aponta o que considera os dois maiores problemas do uso de patentes para análise econômica, a saber: o de classificação e o de “variabilidade intrínseca”. O primeiro, considerado de ordem técnica, refere-se ao problema de como alocar os dados de patentes, organizados por firmas ou classes tecnológicas, em indústrias economicamente relevantes ou agrupamentos de produtos, por exemplo. Dado que o sistema de classificação (classes e subclasses) está baseado primordialmente em princípios tecnológicos e funcionais e, muito raramente, relacionado à noções econômicas de produtos e indústrias bem definidas. Griliches (1990) sugere que a pergunta a ser feita aos dados definirá, em última instância, sua classificação em cada pesquisa.

Antes de tentar qualquer classificação, é preciso enfrentar a ambiguidade inerente da tarefa. Queremos atribuir a invenção à indústria em que foi fabricada ("origem"), à indústria que provavelmente a produzirá, ou à indústria que usará o produto ou processo resultante e cuja produtividade poderá se beneficiar (destino ou indústria de "uso")? [...] Depende de qual pergunta deve ser feita sobre os dados. [...] Essa

90 “[...] a patente does represent a minimal quantum of invention that has passed both the scrutiny of the patent office as to its novelty and the test of the investment of effort and resources by the inventor and his organization into the development of this product or idea, indicating thereby the presence of a non-negligible expectation as to its ultimate utility and marketability. One recognizes, of course, the presence of a whole host of problems: Not all inventions are patentable, not all inventions are patented, and the inventions that are patented differ greatly in "quality," in the magnitude of inventive output associated with them” (GRILICHES, 1990, p. 1669).

diferença de perguntas se reflete também em diferentes estratégias de classificação adotadas por diferentes pesquisadores (GRILICHES, 1990, p.1666)⁹¹.

O segundo problema deriva do fato de as patentes diferirem no que tange a significância técnica e econômica. Dentre as muitas abordagens para tratar essa questão, como pesquisas com titulares de patentes sobre retornos auferidos e potenciais valores de mercado de seus direitos de propriedades, avaliação do valor implícito na decisão sobre pagar a taxa de renovação da patente, análise econométrica da relação entre número de patentes e outras variáveis denominadoras de valor (como lucro e valor de mercado das ações), Griliches (1990) cita o crescente interesse no uso da contagem de citações de patentes como indicadores alternativos de diferencial de qualidade.⁹²

Quando se trata especificamente do uso de dados de citações, surge a questão do problema de truncagem. Ou seja, geralmente, patentes mais antigas possuem um número maior de citações por patentes subjacentes, quando comparadas às patentes mais novas. Portanto, no ato de escolha e recorte de um determinado período de tempo para uso dos dados, cria-se o problema de truncamento, aprofundando, para patentes mais recentes. Trajtenberg (2002) apresenta uma das primeiras tentativas de contornar esse problema testando idade *versus* importância da patente, na busca de e tentando inferir a extensão do viés na contagem de citações advinda da truncagem inevitável dos dados.

3.1.1. Da literatura sobre uso de patentes e citações: uma breve síntese.

A origem da análise quantitativa da mudança tecnológica se localiza no período imediato pós II Guerra Mundial, tendo na obra organizada por Richard Nelson (1962), *The rate and direction of inventive activity*, o primeiro esforço de síntese e guia das práticas em voga. A obra sumariza as primeiras linhas de pesquisa em análises empíricas do processo inventivo e inovativo e constituiu, por muitos anos, um ponto de referência. Mais especificamente, contém o artigo de Simon Kuznets que evidencia a dificuldade em mensurar os resultados do processo inventivo, discutindo os problemas de se definir e mensurar a magnitude das invenções, a relação entre o

91 Tradução do original em inglês.

92 "It should be noted here that patent citations differ from usual scientific citations to the work of others in that they are largely the contribution of patent examiners whose task is to delimit the reach of the new patent and note the context in which it is granted. In that sense, the "objectivity" of such citations is greater and may contribute to the validity of citation counts as indexes of relative importance." (GRILICHES, 1990, p.1689)

significado tecnológico e econômico da invenção, a distinção entre o custo de se produzir uma invenção e o valor que esta cria, e as consequências de uma distribuição altamente enviesada do valor das invenções. Especificamente sobre patentes, Kuznets (1962) considerou os benefícios e as desvantagens das estatísticas de patentes e solicitou que se fosse além da mera contagem de patentes, utilizando as ricas e detalhadas informações sobre o processo inventivo em si, reveladas nos documentos de patentes.

Já em 1966, Jacob Schmookler, em *Invention and Economic Growth*, apresenta o resultado de um processo de análise metódica dos registros de patentes para compilar centenas de séries temporais de totais de patentes por setor. Em especial, deu atenção às questões metodológicas surgidas do uso desses dados, a saber: a dificuldade em se identificar patentes com indústrias a partir da classificação tecnológica que lhe é atribuída pelo escritório de patente. Ademais, argumentou que as estatísticas de patentes representam uma fonte única de informações sistemáticas sobre o processo inventivo.

O volume *R&D, Patents and Productivity* (1984), organizado por Zvi Griliches, é constituído por artigos apresentados em uma conferência do *National Bureau of Economic Research* (NBER) que, em sua maioria, se baseiam na junção de bases de dados de patentes (USPTO) com microdados sobre os gastos com P&D realizados pelas firmas americanas. Ou seja, apresentam a compilação de um amplo conjunto de dados em painel ao nível da firma que combinava o total de patentes com P&D e outras informações financeiras. A partir desse esforço, várias pesquisas foram realizadas com base nesse conjunto de dados estabelecendo muito do que agora tem-se como “fatos estilizados” sobre P&D e patentes ao nível da firma (JAFFE E TRAJTENBERG, 2002). Alguns dos principais resultados incluem a constatação de que, na análise *cross-section*, patentes são aproximadamente proporcionais a P&D, com a proporção variando por indústria e sendo maior para pequenas firmas; na análise em painel, tem-se que, para algumas firmas, patentes são correlacionadas com P&D com retornos decrescentes; nos modelos multivariados incluindo P&D, patentes e medidas de performance (ex.: crescimento da produtividade, lucratividade, valor de mercado), a maior parte da informação resultante é sobre a existência de correlação entre P&D e performance. Ademais, a contagem de patentes está mais fracamente correlacionada com performance, e geralmente não possuem poder de explicação incremental uma vez que o P&D é incluído. E, por fim, mostrou-se que informações detalhadas sobre a composição tecnológica das patentes das firmas podem ser usadas para localizar os programas de pesquisa das firmas no “espaço tecnológico”, e que

variações ao longo desse espaço nas oportunidades tecnológicas e “spillovers” de P&D tem efeitos mensuráveis sobre a performance da pesquisa.

Os anos 1970 e 1980 apresentam muitos desenvolvimentos conceituais na modelagem do processo de pesquisa e do papel das patentes nesse processo: Griliches (1979) e Griliches e Pakes (1984) estendem e refinam o conceito de “função de produção do conhecimento”, uma relação estocástica na qual o atual investimento em P&D, o estoque de conhecimento já existente na firma, e o conhecimento de outras fontes combinam-se para produzir novo conhecimento. Pedidos de patentes podem ser vistos como um indicador do sucesso desse processo estocástico de produção de conhecimento, com a “propensão a patentear” – a proporção de patentes para a produção de conhecimento não observada – possivelmente variando ao longo do tempo e instituições. Griliches (1979) também sugeriu que a possibilidade de um excedente de retorno social na pesquisa deveria ser explicitamente modelada em relação aos fluxos de conhecimento entre e ao longo dos diferentes agentes econômicos.

Já o trabalho de Schankerman e Pakes (1985, 1986), usando informações sobre taxas pagas pela renovação de patentes nos países europeus, permitiu estimar a distribuição dos valores das patentes (privadas), induzidas pelas frequências de renovação e pela magnitude das taxas de renovação em todas as etapas. Essa linha de pesquisa forneceu sólidas evidências empíricas sobre a extensão da heterogeneidade nos valores das patentes e os estímulos para pesquisas adicionais usando aspectos novos dos dados das patentes (PAKES E SIMPSON, 1989).

No artigo de 1990, Griliches argumenta a favor do uso de estatísticas de patentes, por estarem disponíveis e relacionadas à inventividade, por definição, dada a ausência de boas medidas de todo o processo pela qual perpassa a atividade inventiva. Advoga em favor do uso de indicadores de patentes apesar das dificuldades que possam surgir de seu uso e interpretação. Para tanto, descreve a natureza das patentes e os tipos de dados gerados pela sua emissão, disponibilidade dos dados, e os principais problemas que podem surgir da sua utilização em análises econômicas. Ademais, realiza uma revisão abrangente dos trabalhos relativos ao uso de patentes para estudo empíricos antes de 1990. Focando, em particular no trabalho de Jacob Schmookler e nas questões levantadas por ele – em especial no que tange a capacidade de medição das estatísticas de patentes -, assim como nos trabalhos realizados pelos grupos da NBER, Scherer, Yale e SPRU. Segundo Griliches, sob todo os trabalhos pairam as seguintes questões: as estatísticas de patentes podem ser usadas para o quê? Podem ser utilizadas para

interpretar tendências de longo prazo? Ele argumenta que, dada a quantidade de informações contidas no documento de patentes, deve-se ir além do que a simples contagem agregada para um determinado ano ou firma. Dentro das possibilidades estariam o estudo da distribuição geográfica de uma determinada invenção e a investigação de redes e padrões de citações.

O volume organizado por Jaffe e Trajtenberg (2002) é uma consequência direta e uma resposta a essa trajetória de pesquisa, segundo os próprios organizadores. Compilam um conjunto de dados altamente detalhado sobre todas as patentes concedidas nos EUA entre os anos de 1963 e 1999, bem como as citações que elas fizeram. Os dados compreendem uma ordem de, aproximadamente, 3 milhões de patentes e 16 milhões de citações. Os artigos, constantes no livro organizado por Jaffe e Trajtenberg (2002), usaram os dados de patentes para alcançar um tratamento empírico da quantificação da importância ou valor das inovações, na medição dos fluxos de conhecimento tecnológico, e na caracterização dos desenvolvimentos tecnológicos e impactos de instituições ou países particularmente. Publicam estatísticas de patentes atualizadas e expandidas para utilização por outros pesquisadores. E analisaram as questões metodológicas e de interpretação que surgem do uso de estatísticas de patentes. Ademais, desenvolveram algumas métricas baseadas em citações destinadas a capturar diferentes aspectos das inovações patenteadas, tais como “generalidade”, “originalidade” e *time lags* (tempo de defasagem) das citações.

Desenvolveram o uso de citações de patentes para rastrear fluxos de conhecimento tecnológico de um inventor para outro e a utilização do número e o caráter das citações, finalmente recebidas por uma determinada patente, para caracterizar o impacto tecnológico e econômico de uma determinada invenção. Fornecendo, assim, uma maneira empiricamente significativa de examinar a questão da magnitude das invenções. Essa abordagem também fornece uma maneira de lidar com o aparentemente "baixo conteúdo de informações" das contagens de patentes encontradas em grande parte do trabalho econométrico da década de 1980: ponderar as patentes pelo número de citações que elas recebem posteriormente produz uma medida muito mais significativa da produção inventiva do que simples contagens de patentes. Também forneceram uma análise de como o processo que gera os dados afeta sua interpretação e apresentando técnicas econométricas para lidar com algumas dessas questões.

Um dos primeiros trabalhos a usar sistematicamente dados de citações de patentes, e não apenas a simples contagem de patentes, nesse caso para uma área específica, foi o trabalho de

Trajtenberg (2002) na busca de captar a importância e o valor das inovações relacionadas à Tomografia Computadorizada. Mostrou-se que a associação de indicadores construídos a partir das citações com outras medidas independentes – como valor de mercado – revelou-se significativo na captura dos ganhos sociais desta tecnologia. Mais especificamente, a contagem de patentes ponderada pelas citações estava positivamente correlacionada aos ganhos sociais das inovações em C&T. “Assim, a frequência de citações subsequentes mostrou-se um indicativo da “importância” das inovações subjacentes, como medido pelos ganhos de bem-estar resultantes calculados com base em um modelo discreto de demanda por escolha” (TRAJTENBERG, 2002, p. 9)⁹³. A hipótese central é a de que citações de patentes podem informar o valor econômico das inovações e não apenas ser indicativo da importância tecnológica.

Enquanto a contagem simples de patentes tem se mostrado como um bom indicativo somente do lado da oferta, como refletido pelos gastos com P&D (GRILICHES, 1984), observa-se que patentes variam enormemente em sua significância tecnológica e econômica, o que torna sua simples contagem pouco significativa para além do lado da oferta do processo inovativo. No entanto, as citações podem refletir um tipo de relação causal entre as patentes citantes e as citadas, o que é consistente com a visão da inovação como um processo contínuo e incremental, pontuado por ocasionais avanços. Observou-se, também, que as citações por patentes decrescem drasticamente ao longo do tempo, mas ainda podem ser bons indicadores do valor da inovação. Trajtenberg (2002) observa que a pura contagem atribui o mesmo valor para todas as patentes e sugere o uso de citações de patentes como índice de valor ou importância das patentes – sugestão inspirada nos estudos bibliométricos do uso de citações de publicações científicas –, por exemplo, contar o número de vezes que cada patente foi citada em patentes subsequentes e usar o número para computar contagens de patentes ponderadas.

A potencial significância das citações de patentes pode ser inferida do fato de que durante o processo de exame do pedido de patente, é verificado se a invenção reivindicada não foi protegida antes ou se parte dela não está protegida por patentes anteriores, limitando o escopo da proteção requerida. Se são encontradas patentes relacionadas, elas são citadas pelos examinadores no documento da patente concedida. Assim, o número de vezes que um documento de patente é citado, pode ser uma medida de seu significado tecnológico. Além do

93 Tradução do original em inglês.

mais, existe uma dimensão legal nas citações de patentes, uma vez que elas representam uma limitação no escopo dos direitos de propriedades estabelecidos pelas reivindicações de uma patente, o que tem peso em juízo. Igualmente importante, o processo de se chegar à lista final de referências, que envolve o solicitante e o examinador, parece gerar os incentivos corretos para que todas as patentes relevantes sejam citadas e apenas elas. Esses fatos dão bons fundamentos para a pressuposição de que a contagem de citações podem ser bons indicativos da importância tecnológica da patente.

Importante lembrar que a importância tecnológica tem sido pensada apenas como relacionado ao lado da oferta das inovações. Outra questão é se a contagem de patentes pode ser indicativa do valor (*ex post*) das inovações descritas nas patentes citadas: a própria existência das patentes subsequentes atesta o fato de que as patentes citadas abriram o caminho para uma linha de inovação tecnologicamente de sucesso. Ademais também atesta o sucesso econômico, dado que as patentes subsequentes são o resultado de esforços inovativos custosos assumidos por agentes em busca de lucro. Se as citações continuam a aparecer para uma determinada patente, a invenção iniciada na patente citada provou-se ser valiosa. Assim, Trajtenberg (2002) conclui que o uso das citações como indicadores estão associadas a uma particular visão do processo inovativo, na qual a inovação é vista como um processo contínuo no tempo que possui uma natureza predominantemente incremental e cumulativa, pontuada por ocasionais avanços que trazem esforços inovativos subsequentes e os direciona dentro de novos canais

Trajtenberg, Henderson e Jaffe (2002) exploram o uso das citações como indicadores do transbordamento (*spillovers*) de conhecimento e, para tanto, constroem um conjunto de dados constituídos de patentes concedidas às universidades em 1975 e 1980, e uma amostra similar para corporações. Desenvolvem novas medidas baseadas em citações e exploram quão informativas elas são da natureza variada das invenções. Se baseiam na hipótese de que, em média, as patentes de universidades tendem a ser mais fundamentalmente tecnológicas, dado que estas instituições desenvolvem mais pesquisa básica. Encontram, de fato, que as medidas, baseadas em citações, relacionadas à “basicidade” das invenções obtém escores mais altos para as patentes de universidades do que para as das corporações. Testam as hipóteses conjuntas de que, i) citações são uma proxy para o impacto tecnológico, ii) invenções de universidades possuem maior impacto tecnológico. Uma questão que surge de singular importância é: mesmo que as citações sejam indicativas de importância, em si mesmas não possuem nenhum mecanismo natural de calibração: ou seja, não há como dizer a priori quantas citações são

“suficientes” ou se configuram um grande número. O que é particularmente verdadeiro quando existe o problema de truncagem, fazendo-se necessário o uso de grupo de controle ou referência.

Os pesquisadores apresentaram algumas medidas construídas do caráter e significância tecnológica das inovações: “generalidade” é definida como uma medida da extensão na qual as citações recebidas por uma patente estão amplamente dispersas ao longo das classes tecnológicas – mantendo constante o número de citações, os autores sugerem que quanto maior a dispersão das citações ao longo das tecnologias, mais amplo o impacto tecnológico, e maiores os potenciais retornos sociais. “Originalidade” é uma medida estatística análoga, porém, construída com base nas citações feitas ao invés das recebidas – é pressuposto a noção de que pesquisas “originais” tendem a ser sintéticas, com base em pesquisas anteriores de vários campos diferentes. Também introduzem as medidas de taxas de “auto-citação” (proporção de citações feitas pelo mesmo titular a quem pertence a patente citada) e conjecturam que tais taxas podem refletir o grau de apropriação dos potenciais “*spillovers*” de uma dada invenção pela organização proprietária. Outra medida também apresentada se relaciona à extensão da dependência de ciência versus tecnologia, como mensurada pela proporção de citações à literatura científica (“citações não-patentes”) para todas as citações de patentes e não patentes que aparecem na patente.

Caballero e Jaffe (2002) introduzem inovações metodológicas através de uma “função de citação” que modela o processo de geração de citações com os efeitos combinados da difusão obsolescência gradual. O primeiro efeito faz com que as taxas de citação subam à medida que o tempo passa após uma invenção, enquanto a obsolescência gradual faz com que as taxas de citação diminuam com o tempo. Também exploram como várias observações sobre coortes de patentes citantes e citadas, ao longo do tempo e em diferentes campos tecnológicos, podem ser usadas para identificar empiricamente até que ponto as taxas de patentes observadas e as taxas de citação são afetadas por variações na propensão a patentear e na propensão em fazer citações de patentes.

Jaffe, Trajtenberg e Henderson (2002) demonstram que a localização geográfica de transbordamento de conhecimento é estatisticamente significativa quando capturada por citações de patentes. Encontram que patentes citantes tendem a se localizar nas mesmas áreas metropolitana, cidade ou país da patente citada. No entanto, há evidências de que as diferenças geográficas diminuem nas taxas de citações com o passar do tempo. Esse resultado corresponde

com a intuição e corrobora o teste: qualquer que seja a vantagem inicial da proximidade geográfica em termos de transmissão de conhecimento e estímulo para criação adicional de conhecimento, a própria natureza etérea do conhecimento define a diminuição dessa taxa ao longo do tempo. Outras dimensões de “proximidade” entre invenções – tecnológica, institucional, etc. - também se mostram relevantes. O trabalho ajuda a estabelecer que a noção de transbordamento de conhecimento pode ser rastreada empiricamente geograficamente e ao longo de outras dimensões, e que a junção entre geografia e tempo importa.

Jaffe e Trajtenberg (2002) examinaram em mais detalhes os padrões de localização geográfica e quantificaram diferenças nos fluxos de transbordamento entre países, descobrindo algumas idiosincrasias como, por exemplo, a tendência dos inventores japoneses em se basear em inovações mais recentes quando comparados aos inventores de outros países. Também demonstraram o potencial de pesquisa do uso da função de citação para o estudo da rede complexa de transbordamentos de conhecimento à medida que eles fluem ao longo das localizações, tecnologias e do tempo.

Por sua vez, Henderson, Jaffe e Trajtenberg (2002) analisam as mudanças nos patenteamentos das universidades após o Bayh-Dole Act, em 1980, que pretendia aumentar a comercialização das invenções derivadas das universidades e financiadas com recursos públicos. De fato, nota-se um aumento expressivo de patentes concedidas às universidades nas décadas de 1980-90. No entanto, as medidas, baseadas em citações, de importância e generalidade, mostraram que, na média, a significância das invenções das universidades diminuíram após 1980. Indicando que a mudança no regime político promovido pela Lei aumentou a quantidade de transferência de tecnologia pelas universidades, mas não sua qualidade.

Até aqui tratou-se de como captar a dinâmica do conhecimento tecnológico através o uso de dados de patentes, a partir do que elas significam enquanto indicadores do resultado do processo inovativo, especificamente quando utilizadas as informações contidas em seu documento, com especial atenção às citações feitas à outras patentes. Seguida da apresentação de uma breve revisão da literatura que trata do uso desses dados para a investigação empírica da mudança técnica. No entanto, outra dimensão do processo inovativo se traduz na produção de conhecimento científico, especialmente em se tratando de revoluções tecnológicas que, em geral, testemunham um maior dinamismo na produção científica. Portanto, a próxima seção se

ocupará das formas de se captar tal dinâmica do conhecimento científico e de uma breve revisão metodológica.

3.2. Artigos como indicadores da dinâmica científica: uma revisão conceitual e metodológica.

Dada a definição das características das *general purpose technologies*, bem como dos paradigmas tecnológicos, discutidas anteriormente no Capítulo I, fica claro o importante papel desempenhado pela ciência na evolução de qualquer GPT, proporcionando o desenvolvimento dos cluster de tecnologias e suas aplicações, e na composição da base de conhecimento sob a qual o paradigma tecnológico se desenvolve. Por outro lado, existe também uma relação causal inversa, onde as descobertas científicas, por sua vez, são afetadas pelos avanços na tecnologia (STEPHAN, 2010). Ou seja, a tecnologia influencia a atividade científica de maneiras numerosas e difusas. Rosenberg (2006) enfatiza a reconhecida influência de certas preocupações tecnológicas no crescimento do conhecimento científico. E ainda, a possibilidade de um conhecimento científico de grande generalidade originar-se de um problema particular num contexto limitado. A própria difusão da internet a partir dos anos 1990, por exemplo, aumentou a velocidade de conexão e diminuiu os custos de colaboração entre instituições (STEPHAN, 2010). Impulsionando a colaboração na produção científica e permitindo uma maior interação entre agentes e instituições, bem como de áreas distintas, potencializando a interdisciplinariedade. A internet mudou a forma de comunicação científica. A própria WWW, a princípio, foi pensada para aumentar a troca de conhecimento entre pesquisadores. Isso corrobora o fato de a ciência possuir impacto econômico, sendo fonte de crescimento (STEPHAN, 2010).

O fato de o relacionamento não ser linear ter entrado mais recentemente na sabedoria convencional, já que o papel da tecnologia na formação dos avanços científicos foi investigado e articulado por Rosenberg e Mokyr, além de outros. A não linearidade refere-se não apenas ao papel desempenhado pelos equipamentos nas descobertas científicas, mas também ao papel que as inovações tecnológicas desempenharam na promoção de insights científicos, bem como no seu papel de incentivar os cientistas do setor público a desenvolver novos programas e agendas de pesquisa (STEPHAN, 2010, P.259)⁹⁴.

94 Tradução do original em inglês.

Dada a importância da ciência, se faz necessário entender as especificidades do conhecimento científico e da sua produção, algo do qual a Economia da Ciência vem se ocupando. O conhecimento científico possui uma natureza pública e, portanto, um potencial de difusão elevado. O conhecimento não se esgota quando compartilhado, e, uma vez tornado público, não se pode excluir outros de sua utilização. Além do mais, o uso extensivo do estoque de conhecimento não o diminui, pelo contrário, o aumenta (STEPHAN, 2010). O que possibilita uma capacidade de expansão infinita. Portanto, é importante entender a estrutura do sistema de recompensas usado para encorajar a produção e o compartilhamento deste conhecimento. Sabe-se que mercados competitivos falham em prover bens públicos por conta de suas características e em fornecer os incentivos corretos para sua produção. Então, o resultado é a presença de uma estrutura de recompensa não mercadológica, que consiste no reconhecimento de que a prioridade na descoberta científica é um tipo de direito de propriedade. O objetivo do cientista, de acordo com tal princípio, seria o de estabelecer a prioridade da descoberta, sendo o primeiro a comunicar um avanço no conhecimento. Assim, as recompensas de prioridade são o reconhecimento concedido pela comunidade científica por ser o primeiro (STEPHAN, 2010).

O reconhecimento de prioridade pode assumir formas variadas dependendo da importância que a comunidade científica atribui à descoberta em questão. Têm-se a prática de atribuir o nome do cientista à descoberta, o reconhecimento em forma de prêmios que concedem prestígio (e uma premiação em dinheiro) e as publicações. Segundo Stephan (2010), publicações configuram a menor forma de reconhecimento, porém, um passo necessário para estabelecer prioridade. Ademais, a forma comum de mensurar a importância da contribuição de um cientista é contar o número de citações recebidas por um artigo de sua autoria, ou o número total de citações recebidas pela integralidade de seu trabalho. Stephan (2010) e van Raan (2019) atribuem a evolução de mecanismo de contagem de citações à importância que essa prática possui até os dias atuais. Além do mais, boa parte da remuneração financeira dos cientistas - salários e recursos - é baseada em produtividade, muitas vezes avaliada através da contagem de publicações e citações. O que configura uma situação na qual os cientistas, potencialmente, assumem a maior parte do risco da atividade científica.

É importante descrever os tipos de comportamento que essa estrutura de incentivos provoca, como a necessidade de publicar rápido, no entanto, sabe-se da existência de diferenças de tempo de publicação entre as diferentes disciplinas do conhecimento. E muita energia gasta em

estabelecer prioridade sobre o rival. Uma característica da ciência que contribui para a criação de tal estrutura de recompensas é o fato de não ser possível monitorar o esforço científico.

Uma das características do processo no qual o conhecimento é produzido é que a ciência leva tempo e que o conhecimento desempenha um papel de crucial importância nas descobertas, pois, intensifica a “corrida”/competição, dado que muitos pesquisadores tem acesso à base de conhecimento para resolver o problema. Além do mais, o conhecimento tanto pode estar incorporado nos cientistas que trabalham na pesquisa ou não incorporado, porém disponível na literatura, e os diferentes tipos de pesquisa podem estar mais relacionados a um desses tipos. No entanto, a base de conhecimento de um cientista pode se tornar obsoleta se este falha em acompanhar as mudanças na disciplina e o fato de haver evidências de que conhecimento demais pode prejudicar a descoberta dado que entrava o pesquisador. A produção de conhecimento também requer recursos que variam em complexidade, custo, volume e especificidades. A produção científica, dessa forma, pode depender dos graus de sofisticação das ferramentas de pesquisa, o que pode, em última instância, determinar o ritmo do avanço científico.

A importância do equipamento é uma das razões para enfatizar a não linearidade das descobertas científicas. A pesquisa científica pode levar ao avanço tecnológico, mas a tecnologia afeta muito os avanços da ciência. A história da ciência é a história de quão importantes são os recursos e equipamentos para a descoberta (STEPHAN, 2010, p.229)⁹⁵.

O tipo de regime de financiamento da ciência influencia diretamente na velocidade do avanço do conhecimento. Ademais a pesquisa científica pode alcançar resultados não previstos, ou que não se constituíam o objetivo inicial da pesquisa, mas de especial relevância em termos de descoberta científica. A descoberta científica possui uma natureza imprevisível.

Assim, talvez seja mais apropriado pensar em acaso como o ato de encontrar respostas para perguntas ainda não colocadas. Importantes avanços médicos, por exemplo, vieram de pesquisas fundamentais e não dirigidas a missões. Um cientista que estuda caracóis marinhos encontrou uma nova e poderosa droga para a dor crônica. Um medicamento contra o câncer amplamente utilizado surgiu de estudos sobre como a eletricidade afeta os micróbios (STEPHAN, 2010, p.232)⁹⁶.

95 Tradução do original em inglês.

96 Tradução do original em inglês.

Patentear ou publicar podem configurar atividades complementares. Patentes podem ser um resultado lógico de atividades de pesquisa que foram inicialmente desenhadas, principalmente, com objetivo de publicação. Estudos revelam que, em algumas áreas do conhecimento, patentear possui efeitos positivos sobre publicação. No entanto, direitos de prioridade e de propriedade diferem quanto aos incentivos para divulgação das descobertas na pesquisa de forma mais rápida. Os direitos de prioridade exigem a rápida divulgação para que seja estabelecido, já os direitos de propriedade tentam garantir retorno econômico do uso da tecnologia, o que pode desencorajar a rápida divulgação do conhecimento, dado que algumas formas requerem o reconhecimento de sua natureza pública – como as patentes.

A distinção é tão crucial que Dasgupta e David (1987, p. 528) argumentam que os dois tipos de direitos de propriedade e as implicações que eles têm para a apropriabilidade e a divulgação diferenciam a ciência da tecnologia: “Se alguém se junta ao clube de ciências, suas descobertas e as invenções devem ser completamente divulgadas, enquanto no clube de tecnologia essas descobertas não devem ser totalmente reveladas ao restante dos membros” (STEPHAN, 2010, p.249)⁹⁷.

3.2.1. Revisão de estudo bibliométricos para análise da dinâmica e evolução do conhecimento científico;

Com base na discussão anterior, define-se então, que o conhecimento científico é o resultado da pesquisa científica (STEPHAN, 2010; DASGUPTA e DAVID, 1994). Ou ainda, como sugere Van Raan (2014), o progresso científico pode ser definido com o aumento substancial do nosso conhecimento sobretudo, ou do estoque de conhecimento disponível (NELSON, 1959). E que artigos, publicados em periódicos e anais de conferências, são um dos maiores produtos da atividade científica⁹⁸, possibilitando o surgimento de estudos quantitativos sobre a ciência e sua dinâmica, através de métodos estatísticos e matemáticos para analisar e construir indicadores sobre a produção científica. As publicações oferecem diversos elementos para "medir" aspectos importantes da ciência, como os nomes de autores, endereços institucionais, periódico (que indica o campo de pesquisa e o status), referências (citações), conceitos (palavras-chave e combinações de palavras-chave), resumos e títulos.

97 Tradução do original em inglês.

98 Scientists have communicated (and codified) their findings in a relatively orderly, well defined way since the 17th century. Particularly is the phenomenon of serial literature crucial: publications in international journals. Thus communication, i.e., exchange of research results, is a crucial aspect of the scientific endeavour. Publications are not the only, but certainly very important elements, in this process of knowledge exchange (VAN RAAN, 2014, p.25).

Van Raan (2019) define a cientometria como o estudo quantitativo da ciência com o objetivo de avançar nosso conhecimento sobre o desenvolvimento da ciência e sua estrutura de comunicação, bem como em relação aos seus aspectos sociais, tecnológicos e socioeconômicos. A bibliometria, por sua vez, um ramo dentro da cientometria, se dedica à pesquisa da comunicação científica através de dados de publicação, citações e periódicos (BOYACK E KLAVANS; 2019; VAN RAAN, 2014 e 2019). Considera-se a criação do *Science Citation Index*, por Eugene Garfield, um grande avanço que possibilitou a análise estatística da literatura científica em grande escala, marcando a ascensão da bibliometria como um campo poderoso dentro dos estudos da ciência (VAN RAAN, 2014).

Os cientistas são fascinados por características básicas como simplicidade, simetria, harmonia e ordem. O Science Citation Index permitiu que De Solla Price começasse com o desenvolvimento de uma "abordagem física" da ciência, na qual ele tentava encontrar leis para prever novos desenvolvimentos, inspirado nas ideias da mecânica newtoniana e estatística. Nesta perspectiva, as medidas quantitativas da ciência, "indicadores", são guias para encontrar e, como uma próxima etapa crucial, para compreender tais características básicas. A característica mais básica diz respeito à dimensão cognitiva: o desenvolvimento do conteúdo e da estrutura da ciência. Muito dos indicadores científicos na superfície mundana se relacionam com a dimensão social da ciência, em particular com aspectos formulados em questões como "Quantos pesquisadores? Quanto dinheiro é gasto em ciência? Quão 'bons' são os grupos de pesquisa? Como funciona a comunicação na ciência, particularmente qual é o papel dos livros, periódicos, conferências (Borgman, 1990)? E há mais tempo do que muitas vezes percebemos, há outra questão: "Qual é o lucro econômico das atividades científicas?" (VAN RAAN, 2014, p. 20)⁹⁹.

Em 1978 foi fundado o periódico *Scientometrics*, a principal publicação de estudos quantitativos da ciência, marcando a emancipação dessa área de pesquisa. No entanto, Van Raan (2014) salienta a ausência de uma teoria final da ciência que forneça a metodologia para mensurar ciência, ou seja, uma mecânica clássica da cientometria que guie a construção e uso dos indicadores de uma forma definida:

[...] ainda não podemos desenvolver um modelo teórico sólido da 'sociologia do conhecimento', pois simplesmente precisamos de mais trabalho empírico com base na riqueza de dados disponíveis e futuros, a fim de desenvolver uma melhor compreensão quantitativa dos processos pelos quais a ciência e a sociedade influenciam mutuamente os progressos uns dos outros (VAN RAAN, 2014, p. 21)¹⁰⁰.

99 Tradução do original em inglês.

100 Tradução do original em inglês.

Um ramo dos estudos bibliométricos se concentra na avaliação da performance da pesquisa científica, ou seja, na avaliação do impacto do resultado da pesquisa científica feito por cientistas individuais ou grupos de pesquisas. Partindo da hipótese de que cientistas tendem a publicar nas melhores revistas e periódicos possíveis, em especial, quando se trata de grandes progressos. O desempenho científico refere-se à qualidade da contribuição em termos de aumento de nosso conhecimento (progresso científico), conforme percebido por outros pesquisadores com conhecimento (revisão por pares), quantificado e arquivado por citações. No entanto, os estudos evidenciam que esse processo de citação é sem dúvida um processo complexo podendo não fornecer um monitor ideal do desempenho científico para um pesquisador individualmente, porém, quando aplicado a grupos de pesquisa ao longo de um período extenso de tempo pode fornecer um forte indicador de performance científica (VAN RAAN, 2014 e 2019).

As análises de citações são realizadas a partir das práticas de referências dos pesquisadores. Os motivos para fazer ou não uma referência a um artigo em particular, podem variar consideravelmente. Não há, no entanto, evidência empírica de que esses motivos sejam tão díspares ou aleatoriamente concebidos, fazendo com que o fenômeno da citação perca seu papel como uma medida confiável de impacto. Não obstante, regularmente surgem discussões sobre como os autores escolhem suas referências, muitas vezes, em relação a processos de vantagem cumulativa, como o efeito Matthew, por exemplo, nos quais os autores tendem a citar artigos que já são altamente citados. O que pode ser um dos fatores causais da distribuição reconhecidamente enviesada da ciência. Assim, é comum a utilização da análise bibliométrica para acessar a performance científica como complemento às práticas de *peer review*. Com a construção de indicadores que medem impacto das publicações de instituições ou grupos de pesquisa em específico, com relação a medidas internacionais, número de publicações, tipo de periódico e etc. Bem como indicadores a partir do número de citações por publicação, pelo instituto, pelo grupo, num dado período do tempo, etc.

Outra área de estudo da bibliometria, e de especial interesse para o presente trabalho, se debruça sobre a criação e análise de redes bibliométricas de larga escala, destinadas a analisar as relações entre publicações e citações. Boyack e Klavans (2019) atestam que, nos últimos anos, muitas mudanças ocorreram no que tange a criação e análise de redes bibliométricas que descrevam o panorama da Ciência e Tecnologia. Muitas dessas mudanças se devem aos avanços na capacidade computacional e de disponibilidade de dados, resultado direto da era da tecnologia

da informação¹⁰¹ (VAN RAAN, 2014). Os artigos publicados, em jornais e anais de conferência, possibilitam a criação de variadas bases de dados a partir de informações como títulos, resumos e palavras chaves etc. E, dessa forma, redes, com intenção de elucidar a estrutura da ciência, têm sido criadas a partir de todas essas características: redes de artigos podem ser usadas para dividir a ciência em tópicos e especialidades, redes de periódicos são usadas para identificar campos e disciplinas, rede de autores para identificar comunidades e redes de conceitos são criadas a partir de palavras chaves, palavras e frases contidas nos títulos e nos resumos (BOYACK e KLAVANS, 2019). As redes podem ser usadas para identificar um conhecimento científico (ou um campo) que seja revolucionário, interdisciplinar, translacional e transformador. Em suma, a ciência, como o sistema complexo que é, permite ser representada, em seus vários aspectos, por redes bibliométricas distintas, a depender do tipo de análise que se deseja realizar. Ademais, as redes científicas podem ser locais ou globais, estas compreendendo toda a ciência. As redes locais são muito úteis para fornecer alguns tipos de informações, como a estrutura de tópicos dominante em uma especialidade ou os principais pesquisadores e instituições dentro de uma especialidade, o que é de especial interesse para o presente trabalho.

Redes de citações, responsáveis pela evolução da criação de redes científicas, eram utilizadas para identificar tópicos emergentes e para delinear comunidades de pesquisa. O que também conformou a base motivacional para a maior parte do trabalho de cientometria feita nos últimos 60 anos (BOYACK e KLAVANS, 2019). O uso de dados de citações também foi aplicado para entender a história de descobertas científicas chaves e para inferir o impacto da publicação assumindo que esta é relacionada com a frequência em que a publicação é citada (VAN RAAN, 2019). Outros desenvolvimentos se basearam na criação de indicadores recursivos, onde as citações recebem um peso maior quanto mais as fontes citantes são, por sua vez, citadas. Todos esses estudos evidenciam uma característica de decadência exponencial das citações recebidas pelas publicações em função do tempo. Ao se trabalhar com redes bibliométricas precisa-se definir o que serão os nós e o que serão os links, como em qualquer rede, como ficará claro nas próximas seções. E a definição destes elementos depende da pergunta que norteia a pesquisa.

101 O autor ainda pontua que dada o crescente volume de informações disponíveis em meios digitais na Web, e não somente em periódicos, onde os cientistas podem se basear, um novo ramo da cientometria se debruça sobre a análise desse conteúdo, a *Webometrics*.

Nos últimos anos, devido aos avanços nas tecnologias da informação como melhoras na capacidade computacional e softwares, tornou-se possível fazer uma série temporal de mapas científicos, criados através de redes de publicações e suas citações, uma "cinematografia científica" que permite examinar as mudanças nos clusters ao longo do tempo e investigar a natureza da mudança de temas e especialidades de pesquisa. Em alguns casos, a extrapolação de curto prazo pode ser viável na inferência sobre tendências futuras de *hot topics* (BASSECOULARD e ZITT, 2004; BOYACK e KLAVANS, 2019; VAN RAAN, 2019).

Ademais, a bibliometria constitui uma família de métodos entre muitos outros, como os *surveys* e grupos de expert, capaz de oferecer medidas da conexão entre a ciência e a tecnologia (BASSECOULARD e ZITT, 2004)¹⁰². Medir a conexão entre patentes e publicações ajuda a entender a intensidade, orientação e a fonte da relação entre ciência e tecnologia. Além do mais, pode permitir delinear a vizinhança científica de uma invenção, bem como a possível extensão da relevância tecnológica de uma publicação; a investigação da transferência de conhecimento e potenciais *spillovers*: descrevendo a base de conhecimento de tecnologias particulares, bem como revelar a vizinhança tecnológica de temas científicos ou de fronts de pesquisa. A abordagem bibliométrica se aplica tanto às publicações científicas (artigos) quanto às patentes. Os indicadores de ciência e tecnologia baseados em bibliometria assumem que a publicação é uma boa representação do conteúdo da ciência e que as patentes coletam uma ampla parcelada da informação tecnológica, apesar de estas combinarem várias funções. Essas hipóteses têm sido amplamente discutidas na literatura, como apresentado nas seções anteriores, ambas as ferramentas se constituem proxies, não sendo completamente representativas, apresentando distribuições viesadas e outras limitações já discutidas anteriormente. No entanto, tais limitações não invalidam seu uso como indicadores de C&T, principalmente porque os arquivos de publicações e os sistemas de patentes representam uma enorme memória do conhecimento teórico e prático no que tange a cobertura temporal e escopo, sendo acessível, com alto grau de codificação que permite a análise quantitativa¹⁰³.

102 Os autores ainda salientam que o estudo da conexão entre publicações e patentes é somente um caso particular de análises de dados bibliométricos e mapeamento e pode compreender as redes de citações, redes de autores/instituições, esquemas de classificação e redes léxicas.

103 Bassecoulard e Zitt (2004) salientam que transmissão de conhecimento tácito via interações *face to face*, compartilhamento de instrumentos científicos e etc demandam outros métodos de investigação.

No limite, percebe-se que publicações e patentes como fonte de informação possuem características análogas: autor/inventor; instituição/cessionário, referência bibliográfica/referência do sistema de patentes; classificação bibliométrica/classificação oficial; texto completo; referências a literatura científica/referências a patentes e não patentes; etc. O que justifica estender a área de atuação da bibliometria para as patentes, aplicando suas técnicas como a análise de citações, considerando, obviamente as diferenças fundamentais no status e interpretação de cada instrumento.

Em resumo, patentes e publicações, considerados sistemas de avaliação, buscam acessar a novidade/inventividade do resultado requerido pelo autor/inventor. No entanto, como se viu, os regimes de propriedade diferem consideravelmente (regime de sanções legais por exemplo), assim como existem diferentes padrões de divulgação de informação entre publicações científicas e nas patentes. As publicações, em geral, são caracterizadas por apresentar em detalhes dos métodos e resultados das pesquisas científicas para oferecer o maior volume de informação tanto para avaliação e validação por pares (que é a forma de reconhecimento no meio acadêmico) quanto para atrair o maior número de leitores, aplicadores e, principalmente, citações. Por sua vez, as patentes comumente são uma divulgação codificada e com títulos pouco informativos para evitar uma descrição muito precisa da tecnologia que permita o uso pelos competidores. Vale lembrar que a patente é um documento legal assim como um documento técnico, que garante ao aplicante os retornos financeiros da utilização da tecnologia descrita em troca da divulgação compulsória das características técnicas da invenção.

Bassecouard e Zitt (2004) salientam que com a emergência da economia do conhecimento, cada vez mais os grupos de pesquisa, tanto acadêmicos quanto industriais, produzem resultados que poderiam tanto ser publicados em periódicos acadêmicos quanto patenteados. Assim, a escolha da forma de divulgação dependerá dos objetivos de cada grupo. Como visto no capítulo 2, Berners-Lee escolhe por não solicitar o patenteamento da WWW, solicitando a permanência do domínio público da tecnologia, sendo atendido pela CERN. O inventor temia que a cobrança financeira pelo uso da invenção pudesse diminuir ou até mesmo solapar sua difusão, como foi o caso da Gopher, o sistema de gerenciamento de informação da universidade de Stanford.

Dada a discussão subjacente sobre o uso de patentes e artigos como indicadores do conhecimento tecnológico e científico e de suas restrições e limitações, bem como do uso de redes criadas a partir das citações presentes tanto em artigos científicos como nos documentos

de patentes, a próxima seção se ocupará de uma breve revisão da literatura sobre a investigação empírica de tecnologias que se enquadram como GPT, priorizando os estudos que tenham utilizado tais indicadores. O objetivo é apresentar o estado da arte sobre a investigação e detecção de GPTs que sirva de base para a proposta da presente Tese..

3.3. Da literatura sobre a investigação empírica de *General Purpose Technologies*.

Bresnahan (2010) afirma que diferentes estudos tentaram uma investigação econométrica acerca da GPTs, no entanto, a principal dificuldade reside na obtenção de dados sobre a co-invenção nos setores de aplicação. O uso de dados de patentes tem a vantagem de mostrar, via suas citações, os links tecnológicos entre diferentes invenções, o que permite a compreensão não apenas das GPTs, mas também das inovações nos setores de aplicação (BRESNAHAN, 2010, HALL E TRAJTENBERG, 2004).

Hall e Trajtenberg (2004) usaram dados de patentes e suas citações na tentativa de identificar GPTs no período entre 1967 e 2002, e encontram uma lista de patentes que foram citadas ao longo de uma ampla variedade de classes da patentes, demonstrando uma característica de generalidade, ou de serem gerais; estas patentes também eram amplamente citadas, inclusive por patentes por sua vez amplamente citadas, sugerindo um caráter de fecundidade das tecnologias em questão; ademais, esse grupo de patentes pertenciam a classes de patentes que cresceram rapidamente, configurando-se como áreas de importância geral. Outra característica encontrada é a presença de uma longa teia de citações para cada uma dessas patentes. O fato relevante é o grau extremo, ou traço marcante/ forte, de cada uma dessas características neste grupo de patentes, que os autores identificaram como GPT. O método é justamente esse: identificar uma patente que possua traços extremos das características acima mencionadas, o que configuraria tecnologias com estrutura de GPT. As vinte patentes no topo da distribuição das medidas de GPT eram em sua maioria TICs.

Moser e Nicholas (2004) tentaram identificar as características de GPT nas tecnologias relacionadas a eletricidade, tida como uma das tecnologias mais importantes do início do século XX. Para tanto, analisaram as patentes relacionadas à tecnologia de eletricidade na década de 1920, e as citações que estas receberam nas décadas subsequentes. Com o objetivo de examinar

como o conhecimento contido nas patentes de 1920 influenciaram futuras gerações de inovações, usaram as definições de originalidade (uma medida do quão longe, temporalmente, a patente ainda é citada, ou seja, quão original é o conhecimento para estar em uso por longas décadas subsequentes), longevidade (mensurando a velocidade da obsolescência das invenções) e generalidade (que mede a amplitude de aplicação das inovações nos mais diferentes setores). Concluíram que, quando comparadas a outras patentes do mesmo período, relacionadas a outros tipos tecnologias, as patentes investigadas eram mais amplas no escopo e originais. Usaram a definição proposta por Lipsey et al. (1998) sobre os critérios para se identificar uma GPT para analisar se a eletricidade, através dos dados de citações de patente, apresentava uma identificação com tais critérios.

Os trabalhos que foram mencionados acima são estudos de caso de identificação de GPTs para um grupo geral de patentes. O que esta Tese propõe é a investigação da trajetória de uma tecnologia em específico, através da análise do padrão de formação da rede de patentes e de artigos, na busca de identificar semelhanças com as estruturas clássicas de GPTs, como discutido extensamente no capítulo I. Um trabalho com proposta semelhante foi o realizado por Teichert (2017) que propôs uma análise da nanotecnologia como um estudo de caso de GPT, através dos fluxos de conhecimento gerados por esta tecnologia, utilizando redes de artigos científicos e patentes. No entanto, a autora se distanciou da discussão de repercussões paradigmáticas da nanotecnologia, lidando apenas com uma possibilidade de GPT local.

Importante salientar que quando do uso de dados de patentes para identificar GPTs, algumas limitações precisam ser reconhecidas, como o fato de que nem toda inovação é necessariamente patenteada e de que muitas patentes estarem associadas com inovações muito pequenas e pouco importante (Bresnahan, 2010). Hall e Trajtenberg (2004) pontuam que é fundamental, quando da tentativa de identificar patentes com características de GPT, levar em consideração os incentivos para o patenteamento estratégico de tecnologias no período analisado. Bresnahan (2010) ainda aponta que existem spillovers de conhecimento advindas de complementariedades inovativas que, mesmo que patenteadas, podem não possuir citações entre si, dando como exemplo o caso de softwares patenteados que só foram possíveis dado os avanços nos microprocessadores e chips de memórias mais rápidos que, por sua vez, resultam de sofisticadas nas ferramentas de fabricação de circuitos integrados. Em geral, as patentes de softwares não citam as patentes destas ferramentas. Por fim, o autor concluiu que uma tecnologia pode ser geral no sentido de permitir inovações complementares em uma ampla

variedade de setores sem que uma patente daquela tecnologia seja citada por invenções patenteadas em uma ampla gama de setores.

Para a era moderna, os computadores no processamento de dados corporativos são um GPT importante, e muitas das importantes inovações AS associadas a eles são mudanças difíceis de patentear nas práticas de gerenciamento, estruturas organizacionais ou práticas de marketing. Para eletricidade na década de 1920, as inovações complementares importantes foram em melhorias de layout de chão de fábrica difíceis de patentear e invenções de engenharia industrial relacionadas. (BRESNAHAN, 2010, p.781)¹⁰⁴

Essa seção se destinou a uma breve revisão da literatura sobre uso de patentes e artigos, bem como de suas citações, para rastrear e mapear GPTs, na tentativa de estabelecer uma visão do estado da arte no campo da investigação desse tipo de tecnologia. A seção que se segue apresentará uma descrição da criação das duas redes que serão utilizadas na investigação proposta pelo presente trabalho, a saber: a de patentes e a de artigos, compostas por suas respectivas citações. O ponto de partida de ambas as redes, a serem utilizadas, é o artigo de 1992 de Berners-Lee que formaliza a criação da WWW. Em seguida será apresentada a metodologia a ser utilizada na análise das redes, baseada nos desenvolvimentos da ciência de redes para sistemas complexos.

3.4. As redes de patentes e artigos da World Wide Web.

Esta seção visa apresentar os dados de patentes e artigos que serão utilizados nesta Tese para investigar as dinâmicas científicas e tecnológicas da WWW. Para tanto, descreve-se a origem dos dados, qual a base de dados utilizada e como se deu a construção das redes a serem analisadas e manipuladas.

3.4.3. A rede de patentes.

A base de dados utilizada para a construção da rede de patentes foi a PATSTAT, que possui os registros de patentes dos principais escritórios de patentes do mundo, garantindo uma rede significativamente maior de patentes e citações, quando comparado com o uso exclusivo dos registros de patentes do USPTO – *United States Patent and Trademark Office*. Como um dos

104 Tradução do original.

princípios de seleção para constituição da rede, optou-se por utilizar apenas as patentes concedidas, independente do escritório, apesar do ano de contagem da patente ser o de depósito. O PATSTAT Global, publicado pelo *European Patent Office*, possui uma cobertura mundial e contém informações bibliográficas sobre aplicações e publicações, bem como informações legais sobre patentes, tanto de países líderes industriais como de países em desenvolvimento. O PATSTAT Global contém dados bibliográficos relacionados a mais de 100 milhões de documentos de patentes dos principais países industrializados e em desenvolvimento. Ele também inclui os dados de eventos legais de mais de 40 autoridades de patentes contidas nos dados de eventos legais mundiais da EPO (Site EPO, 2020).

A rede de patentes a ser construída no presente trabalho foi pensada na busca de se capturar a relevância da invenção da *World Wide Web* (WWW) através de seu impacto sobre o progresso tecnológico subsequente. Portanto, o ponto de partida foi, naturalmente, a invenção da WWW, formalizada pela publicação de um trabalho científico em 1992 de autoria de Tim Berners-Lee e Cailliau, R., intitulado "*The World-Wide Web*", e apresentado em uma conferência da CERN - *Conseil Européen pour la Recherche Nucléaire*. Tim Berners-Lee inventou a *World Wide Web* (WWW) em 1989, enquanto trabalhava no CERN. A web foi originalmente concebida e desenvolvida para atender à demanda de compartilhamento automatizado de informações entre cientistas de universidades e institutos em todo o mundo, CERN (2020). Então, o ponto de partida é esse trabalho publicado nos anais da conferência da CERN que divulga a invenção de interesse e disponibiliza o conhecimento tecnológico a ser difundido. Ou seja, esse é o primeiro nó da nossa rede. O artigo é o conhecimento que se difundiu facilitando a inovação em outros campos.

Como já mencionado anteriormente, os documentos de patentes possuem dois tipos de citações: à documentos não patentes (bibliografia científica) e às patentes prévias. Dessa forma, o primeiro nível da rede – os primeiros nós - compõe-se por patentes que citaram o trabalho acima descrito, na área do documento de patente relativa às citações não patentes. A hipótese assumida é a de que o conhecimento tecnológico, presente no artigo do Berners-Lee, possui relevância tal que causou a explosão de numerosos progressos tecnológicos posteriores e que esse movimento poderia ser captado pela onda de patenteamento subsequente que citaram o artigo. No segundo nível, o modo de construção da rede muda para as citações feitas às patentes que estão no primeiro nível – ou seja, as que citaram diretamente o artigo do Berners-Lee. No terceiro nível, então, busca-se as patentes que citaram àquelas que compõe o segundo nível e

assim por diante, sempre buscando as citações às patentes constantes do nível anterior. E assim, sucessivamente, os ramos vão sendo construídos. Cada patente na rede constitui um nó a partir do qual nossa rede se ramifica, e cada link nessa rede constitui uma citação.

No primeiro nível, composto por patentes que citaram o artigo do Berners-Lee, no campo de citações não patentes, obteve-se um total de 2.125 patentes que citaram diretamente o artigo. Vale salientar que essas patentes não possuem, necessariamente, a mesma data: o que determina o primeiro nível é quem citou diretamente o artigo, independente do ano de depósito da patente. Assim, patentes que estão no primeiro nível e que são mais recentes em relação ao artigo, possuem, provavelmente, um ramo menor. Dentre as patentes constantes do primeiro nível, encontra-se patente, citando o artigo diretamente, datando do ano de 2010, o que pode ser mais um indicativo da prevalência ou da força dessa invenção inicial.

A rede possui um total de 1.569.977 patentes, sem repetições, e um total de citações de patentes na ordem dos 6.320.366. Ademais, a rede em questão apresenta um total de 65 camadas e, em média, cada patente ou nó, possui 4,5 links, ou seja, cada patente recebeu, em média 4.5 citações de patentes subsequentes. A rede possui patentes concedidas entre os anos de 1992, data da publicação do artigo do Berners-Lee, e 2012, cobrindo um período de 20 anos. A escolha do ano final se deu pelas restrições apresentadas pelos dados disponíveis, como a falta de um número significativo de patentes citando os últimos níveis ou camadas da rede. Ou seja, até 2012 pode-se garantir a coleta de dados confiáveis sem muitos *missing*.

A rede compreende patentes concedidas em variados campos de conhecimento tecnológico, como indicado pelas classes tecnológicas das patentes e demonstrado na próxima sessão, e isso é de especial importância para a captura da difusão tecnológica representada por essa rede. Ou seja, não foi uma rede construída exclusivamente com patentes pertencentes a uma mesma área do conhecimento tecnológico, como tem sido a prática dos trabalhos empíricos sobre a mudança técnica.

A lógica por trás da criação da rede, então, se baseia no fato de que a patente representa uma inovação e dado que seu conteúdo é público, a difusão do conhecimento pode ocorrer. Uma patente pode levar a criação de uma outra patente. Assim, a citação de uma patente por outra significa a potencial difusão do conhecimento original: a difusão do conhecimento. Além de reforçar a visão cumulativa do processo inovativo. No caso específico da rede WWW, o

conhecimento tecnológico inicial a ser difundido teve suas origens no conhecimento científico, no artigo do Berners-Lee. Para entender a dinâmica dessa difusão, voltemos ao processo de formação da rede: inicialmente, o artigo do Berners-Lee (considerado formalmente o inventor da WWW) representaria a inovação tecnológica radical, ou o *big bang* da revolução tecnológica, ou uma GPT, que, por se caracterizar como sendo uma inovação de inerente universalidade, possui potencial para promover uma onda de inovações subsequentes. E isso poderia ser captado pela presença de milhares de patentes subsequentes citando direta e indiretamente tal artigo, nos mais variados campos do conhecimento tecnológico, ou ao longo de todo o espectro de classes tecnológicas, como será evidenciado na próxima seção. Assim se dá o processo de formação da rede.

3.4.4. *A rede de artigos.*

Para a construção da rede de artigos utilizou-se a base de dados de literatura científica *Web of Science* (WoS). A escolha se deu por ser uma base de dados que possui uma cobertura de todos os campos científicos, incluindo referências citadas, possibilitando a construção de uma rede baseada em citação. Ademais, WoS permaneceu como a única base de dados de citações até o ano de 2004 (BOYACK E KLAVANS, 2019).

Redes são em objetos e os links entre estes, comumente chamados de nós e fronteiras, respectivamente. Quando da construção de redes bibliométricas, além da decisão sobre quais dados usar, é necessário definir o que serão os nós e qual característica será usada para criar as fronteiras. Na rede em questão, os nós serão compostos por artigos e os links serão determinados pelas citações apresentadas. O que se deseja analisar é se houve difusão e transformação do conhecimento incorporado no artigo do Berners-Lee (1992), portanto, escolhemos as citações como links entres os artigos, dado que constituem a base de conhecimento sobre a qual um artigo se constrói, como declarado pelo autor de forma intencional.

Dessa forma, a construção da rede a ser analisada se inicia com a definição do artigo de autoria do Berners-Lee (1992), divulgando a criação da WWW, como o primeiro nó da nossa rede – o ano de 1992, então, é o ano inicial de coleta de dados para a rede, terminando no ano de 2015. A partir desse passo, os próximos se assemelham ao caminho seguido quando da construção da rede de patentes. O primeiro nível da rede, ou primeira camada, é composto por artigos que

citaram diretamente, constituindo assim um link direto, o artigo acima mencionado. Indiferente do ano da publicação ou da área de concentração. Estes artigos, citantes do artigo do Berners-Lee, se constituem, então, como os próximos nós a serem analisados. Assim, o próximo passo foi selecionar todos os artigos que, por sua vez, citaram os artigos da primeira camada, ou seja, os artigos que citaram diretamente o artigo denominado como o primeiro nó, ou o nó de origem da rede. A partir da primeira camada, o mesmo procedimento se repete até o ano limite da rede. Os artigos da segunda camada em diante, configuram o que a bibliometria denomina como link indireto entre publicações.

Uma prática comum quando da criação de redes bibliométricas é o calcular a similaridade entre os nós. A abordagem mais simples, denominada citação direta e considerada um link de primeira ordem entre artigos, consiste em tomar cada par de artigo citante e citado como um link, e com mesmo peso cada link. No entanto, existem outras medidas de similaridade mais sofisticadas baseadas na criação de links indiretos entre artigos, como links via cocitação (quando um par de nós é citado por outro nó), acoplamento bibliográfico (um par de nós citando um mesmo outro nó) e acoplamento longitudinal (um nó é citado por outro nó que, por sua vez, é citado por um terceiro nó).

A rede de citação de artigos, resultante do processo descrito anteriormente, entre os anos de 1992 e 2015 somam um total de 13.421.547 artigos e o número de links da rede (ou seja, número de citações entre os artigos) é de 109.288.513.

3.5. A análise de redes: *Science networks*.

Esta sessão se destina a apresentar a metodologia de análise a ser aplicada às redes de patentes e artigos que foram construídas para investigar a existência de uma tecnologia do tipo GPT na WWW. Para tanto, segue uma apresentação sintética da ciência de redes e do método que iremos apresentar, com mais detalhe no tipo de rede que os dados configuram, a saber a *scale free network*.

3.5.1. *Sistemas complexos e ciência de redes.*

Latora et al. (2017) argumenta que “os sistemas sociais, o cérebro humano, a Internet e a World Wide Web são todos exemplos de redes complexas, ou seja, sistemas compostos por um grande número de unidades interconectadas por meio de padrões altamente não triviais de interações” (p.xi)¹⁰⁵. Os autores apontam que o estudo dos sistemas complexos constitui uma ciência relativamente recente acarretando a ausência de uma definição formal e comumente aceita de sistema complexo. No entanto, propõe que sejam entendidos como um sistema formado por um grande número de unidades individuais (indivíduos, componentes ou agentes) interagindo de tal forma que o comportamento do sistema não é uma simples combinação dos comportamentos das unidades individuais fazendo com que o sistema tenha organizações diferentes em escalas de agregação diferentes. Ou seja, é difícil captar o comportamento coletivo a partir do conhecimento dos componentes (BARABÁSI, 2016). Em particular, alguns comportamentos coletivos emergem sem a necessidade de qualquer controle central. Esse tipo de comportamento é o que encontramos nas sociedades humanas em vários níveis, onde as interações de muitos indivíduos dão origem ao surgimento de civilizações, formas urbanas, culturas e economias (LATORA et al., 2017).

Para entender como funciona um sistema complexo é necessário saber como seus componentes interagem entre si, ou seja, será preciso visualizar um mapa das suas diversas conexões, pois é difícil capturar um comportamento coletivo apenas a partir do conhecimento de seus componentes. Essa catalogação dos componentes e suas conexões compõe as redes por trás de cada sistema complexo. Dito de outra forma, cada sistema complexo pode ser representado por uma rede (*network*) que capta sua dinâmica e estrutura, através de nós ou vértices e seus respectivos links ou arestas (BARABÁSI, 2016; MENCZER ET AL, 2020; NEWMAN, 2018). A existência de sistemas complexos nas ciências sociais e humanas deriva de um mundo cada vez mais interconectado, gerando os mais diversos tipos de sistemas, como os tecnológicos, biológicos, econômicos, comunicação, sociais, entre outros. A percepção de que por trás de cada sistema complexo existe uma rede intrincada que codifica as interações entre os componentes do sistema lançaram as bases para o surgimento do que se denomina *Science Network*, uma nova disciplina da ciência que se destina à análise das redes na tentativa de definir regularidades como princípios comuns e leis universais governando a estrutura e a evolução das diversas redes. Pois, as redes definem, em última, instância, a natureza ou o modo de funcionamento dos sistemas a elas associados.

105 Tradução do original em inglês.

Da existência de diversos tipos de sistemas complexos, com distintos processos geradores, tamanho, natureza, escopo, história e evolução, deriva a diversidade entre redes e suas propriedades específicas, como as leis que dão forma e governam seus comportamentos. Entender essas propriedades, codificadas em sua estrutura, ajuda a entender quais fatores contribuem para definir seu comportamento. E, por sua vez, aumenta a compreensão dos sistemas, bem como permite prever e controlar seus eventos. O estudo das redes, em particular nas últimas duas décadas, tem estabelecido que, apesar das múltiplas redes, algumas propriedades universais, governando sua estrutura e comportamento, podem ser observadas. O que contribuiu sobremaneira para o avanço das análises, permitindo a criação de formalizações e ferramentas a serem aplicadas na análise das redes. Percebe-se, também, a possibilidade de sistemas distintos poderem ser representados pelo mesmo diagrama ou rede. Ou seja, a representação em redes oferece uma linguagem comum para estudar sistemas distintos (BARABÁSI, 2016). A ciência de redes objetiva a construção de modelos que reproduzam as propriedades das redes reais.

3.5.2. *Redes: definições*¹⁰⁶.

Como mencionado anteriormente, as redes são diagramas compostos por *nós*, que representam os componentes de um sistema, e por *links*, representando as interações diretas entre os componentes. Definindo os dois parâmetros básicos de qualquer rede: o número de nós (N), sendo o número de componentes do sistema, ou o tamanho da rede; e o número de links (L) sendo o número total de interação entre os nós. Os links entre os nós da rede podem ser direcionados (*directed*) e não direcionados (*undirected*), ou seja, se a interação entre os nós acontece em apenas um sentido (ou direção), são links direcionados, quando os links acontecem em ambas as direções (não importando o sentido) são denominados não direcionados. Diferentes tipos de links resultam em diferentes tipos de redes: redes que possuem apenas links direcionados denominam-se redes direcionadas (*directed networks*); redes não direcionadas (*undirected networks*) contêm apenas links não direcionados, existem ainda aquelas que possuem simultaneamente links de ambos os tipos. Exemplos canônicos, encontrados na literatura de *Science networks*, de redes direcionadas são: a WWW, onde os nós são as *web*

¹⁰⁶ Esta sessão está baseada em Barabási (2016), Newman (2018), Menczer et al. (2020), Latora et al. (2017), Caldarelli e Chessa (2016).

pages que possuem links entre si via URL apontando de um documento para outro; rede de ligações telefônicas, na qual cada celular determina um nó e as ligações feitas como links; e redes de citações, onde cada artigo é representado por um nó e as citações são os links. Como redes não direcionadas tem-se: a Internet, a rede de distribuição elétrica, e rede de colaboração científica, entre outras (BARABÁSI, 2016; MENCZER ET AL. 2020, NEWMAN, 2018).

Outras classificações usuais das redes são: redes ponderadas que possuem links com pesos associados, como uma rede de e-mails onde o peso dos links e a direção representaria o tráfego de comunicação, ou número de mensagens, entre os nós, (MENCZER et al. 2020). Redes bipartidas (*bipartite network*) que são compostas por dois grupos de nós (tipos de nós distintos) nas quais os links somente conectam os nós de grupos diferentes e não os nós do mesmo grupo (exemplos deste tipo de rede são: redes de que capturam a relação entre músicas e artistas, estudantes e classes, produtos e consumidores etc.). Redes *multiplex* (*multiplex network*) possuem múltiplos tipos de links.

Uma importante propriedade dos nós que compõe as diversas redes é o número de links que eles realizam com dos demais, o que é denominado *degree* (ou grau do nó). Em redes não direcionadas o *degree* (k_i) de cada nó se iguala ao número de vizinhos que cada nó possui. Nas redes direcionadas é feita uma distinção entre os links, gerando dois diferentes *degrees*, *in-degree* e *out-degree* (grau de entrada e grau de saída). *In-degree* mede o número de links que cada nó recebe (*inwards links*) – links direcionados ao nó - e *out-degree* contabiliza os links que cada nó realiza (*outwards links*) – ou os links que são direcionados do nó aos demais. O resultante *degree* de um determinado nó será, então, a soma do *in-degree* e *out-degree*. Por exemplo, o *in-degree* de uma página da web é o número de links que outras páginas fazem a ela, enquanto o *out-degree* é o número de links que realiza à outras páginas. Em muitos casos, medidas de centralidade, como o *degree*, são de especial utilidade, pois os nós com os graus mais altos em uma rede, aqueles com mais conexões (denominados *hubs*), também desempenham papéis importantes no funcionamento do sistema e, portanto, o grau pode ser um guia útil para focar nossa atenção nos elementos mais importantes do sistema (NEWMAN, 2018). Outro conceito importante é o de *average degree* (k) - grau médio - da rede, que se resume o número médio de links dos nós da rede. O grau médio para redes não direcionais é dado pela equação abaixo:

$$k = \frac{2L}{N}$$

Para redes direcionais:

$$k = \frac{L}{N}$$

A distribuição de grau (*degree distribution*) possui um papel central na teoria de redes pois muitas propriedades das redes exigem essa medida de distribuição em seu cálculo, ademais, a forma funcional precisa dessa distribuição determina muitos fenômenos da rede, desde sua robustez até a difusão de efeitos (BARABÁSI, 2016). A distribuição do grau (p_k) fornece a probabilidade de que um nó, aleatoriamente selecionado, tenha o grau k . Para uma rede com N nós, a distribuição do grau é dada pela equação que se segue onde N_k é o número de nós que possuem o grau k :

$$p_k = \frac{N_k}{N}$$

É importante ressaltar que muitas redes possuem um pequeno, porém, significativo número de *hubs*, que são nós com uma quantidade elevada de links associados ou grau. Muito foco tem sido colocado na investigação dos efeitos dos *hubs* sobre a performance e o comportamento dos sistemas representados por essas redes e os resultados têm mostrado que os *hubs* possuem um efeito desproporcional, apesar de serem poucos em número ou ocorrerem com pequena frequência, sobre a resiliência da rede e o fenômeno de transporte (NEWMAN, 2018).

A descrição completa de uma rede¹⁰⁷ requer o rastreamento de todos os links presentes, que podem ser representados pela sua matriz de adjacência (*adjacency matrix*). Uma rede direcional contendo N nós terá uma matriz de adjacência com N linhas e N colunas.

$$A_{ij} = \begin{bmatrix} A_{11} & \cdots & A_{1j} \\ \vdots & \ddots & \vdots \\ A_{i1} & \cdots & A_{ij} \end{bmatrix}$$

¹⁰⁷ Redes de citações são redes direcionais, como mencionado anteriormente. Portanto, a partir desse ponto, a apresentação se concentrará na descrição e caracterização de redes direcionais, evitando um texto longo e disperso, para maior objetividade no trabalho.

Onde, $A_{ij} = 1$ se existe um link direcionado do nó j para o nó i ; e $A_{ij} = 0$ se os nós i e j não forem conectados entre si. Ademais pode-se calcular o grau k_i do nó i através da soma entre linha e coluna que fornecem o grau de entrada e saída, respectivamente.

$$k_i = \sum_{j=1}^N A_{ij} \text{ e } k_i^{out} = \sum_{j=1}^N A_{ji}$$

Nas redes reais (*real networks*) o número de nós (N) e links (L) podem variar amplamente resultado em redes de tamanho altamente distintos. Em alguns casos trabalha-se com um subconjunto da rede, na incapacidade de operacionalizar a rede completa. Em uma rede contendo nós $= N$, o número de links pode variar entre $L = 0$ e $L_{m\acute{a}x}$, este sendo o número máximo de links possíveis. $L_{m\acute{a}x}$ é dado por:

$$L_{m\acute{a}x} = \frac{N(N - 1)}{2}$$

Redes que apresentam o número máximo possível de links, $L_{m\acute{a}x}$, são conhecidas como *complete graph* (redes completas), onde cada nó está conectado a todos os outros nós da rede. As redes reais, em geral, apresentam um número de link (L) muito menor do que o $L_{m\acute{a}x}$, definindo sua característica de serem esparsas (BARABÁSI, 2016). Outro conceito importante é o de *path length* (comprimento da trajetória), a trajetória (*path*) é a rota que percorre os links da rede entre um par de nós determinado, e o *path length* representa o número de links que essa trajetória contém. Derivam do *path length* outras propriedades importantes, *shortest path* (a menor trajetória) que determina a trajetória com menor número de links entre os nós i e j . A menor trajetória é denominada distância (d_{ij}) entre os nós i e j , ou apenas d^{108} . A distância máxima ($d_{m\acute{a}x}$) de uma rede, ou a menor trajetória máxima, denomina-se diâmetro da rede (*network diameter*). E o comprimento médio de trajetória (*average path length*) mede a distância média entre todos os pares de nós da rede. A trajetória (*path*) desempenha um papel central na ciência de redes.

108 Em redes direcionais, muitas vezes $d_{ij} \neq d_{ji}$. Além disso, em uma rede direcionada, a existência de um caminho do nó i para o nó j não garante a existência de um caminho de j para i .

Uma funcionalidade chave da maioria das redes é a capacidade de estabelecer uma trajetória entre quaisquer pares de nós, ou assegurar conectividade (*connectedness*). Uma rede será conectada se todos os pares de nós da rede estiverem conectados. Uma rede é desconectada se houver pelo menos um par com $d_{ij} = \infty$. Por sua vez, o coeficiente de agrupamento (*clustering coefficient*) mede o grau de conexão entre os vizinhos de um determinado nó, assim, para um determinado nó i com *degree* k_i o coeficiente de agrupamento local é definido como:

$$C_i = \frac{2L_i}{k_i(k_i - 1)}$$

Onde L_i representa o número de links entre os k_i vizinhos do nó i . C_i é a probabilidade de que dois vizinhos de um nó se liguem e varia entre 0 e 1: $C_i = 0$ se nenhum dos vizinhos do nó i se vincularem; $C_i = 1$ se os vizinhos do nó i formam uma rede completa, ou seja, todos eles conectam-se um ao outro; $C_i = 0,5$ implica que há uma chance de 50% de que dois vizinhos de um nó estejam ligados. Em resumo, C_i mede a densidade local dos links da rede: quanto mais densamente interconectada a vizinhança do nó i , maior é seu coeficiente de agrupamento local. Já o grau de agrupamento da rede como um todo é dado pelo coeficiente de agrupamento médio (*average clustering coefficient*), C representa uma média dos C_i de todos os nós da rede. C é a probabilidade de conexão entre dois vizinhos de um nó aleatoriamente selecionado.

Para caracterizar as redes e poder identificar suas propriedades usualmente adota-se a distribuição de graus (conectividade).

3.5.3. Redes aleatórias (*random networks*).

Evidências empíricas demonstraram que a maior parte das redes reais (*real networks*) não apresentam regularidade e previsibilidade de formato no que tange o número de nós e links, antes sendo altamente aleatórias no seu processo de construção. Portanto, o modelo de redes aleatórias descrita nesta sessão objetiva apresentar as principais características deste tipo de rede na tentativa de reproduzir as propriedades das redes reais que representam.

As redes randômicas (aleatórias) consistem de N nós onde cada par de nós é conectado com probabilidade p , $G(N, p)$ ¹⁰⁹. Em geral, a maioria das redes reais são caracterizadas por conterem alguns nós com um número elevado de links e outros com poucos links ou, até mesmo, com nenhum link. Essa diferença é captada pela distribuição do grau (*degree distribution*), que no modelo de redes reais é representada por uma distribuição de Poisson quando o número de nós (N) é muito maior do que o grau médio (*average degree*) da rede ($\langle k \rangle$)¹¹⁰, ou seja, quando a rede é esparsa. Assim, a distribuição de grau de uma rede aleatória é dada por:

$$p_k = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$$

Essa distribuição possui as seguintes propriedades: apresenta um pico em torno do grau médio ($\langle k \rangle$) e aumentos no grau médio ($\langle k \rangle$) tornam a rede mais densa; a largura da distribuição (dispersão representada por seu desvio padrão) também é controlada pelo grau médio, quanto mais densa for a rede maior é sua largura de distribuição e maior a diferença nos graus apresentados pelos nós. Ademais, a distribuição de Poisson não depende explicitamente do número de nós (N), portanto, prevê que a distribuição de grau de redes de tamanhos diferentes, mas com o mesmo grau médio $\langle k \rangle$, são indistinguíveis entre si.

3.5.4. Scale-free networks¹¹¹.

As evidências empíricas têm mostrado que a maioria das redes reais não são redes randômicas, não sendo bem representadas por uma distribuição de Poisson. As redes reais, em geral, são caracterizadas por apresentarem a maior parcela de seus nós com poucos links e alguns poucos nós com alta quantidade de links, denominados *hubs*, que demonstram desempenhar um papel relevante na determinação do comportamento da rede, não podendo ser assim ignorados. Sabe-se que a distribuição de Poisson subestima a existência desses nós altamente conectados, bem

109 Definição proposta por Gilbert. Existe uma definição alternativa proposta por Erdős and Rényi: $G(N, L)$, onde os N nós são conectados com L links estabelecidos aleatoriamente. Em geral, este modelo é preterido por fixar o número de links.

110 A forma exata da distribuição de grau de uma rede aleatória obedece a uma distribuição binomial, no entanto, quando $N \gg k$, a distribuição de grau é aproximada por uma distribuição de Poisson. Por sua simplicidade analítica, depender de apenas um parâmetro (k) ou *average degree*, é preferida nas análises de redes aleatórias (BARABÁSI, 2016).

111 Vale pontuar que a formalização desse modelo deriva dos estudos da rede formada pela WWW, a partir da percepção de que o modelo aleatório fornecia uma representação pobre para os dados reais.

como aqueles com quantidades pequenas de links, concentrando a distribuição em torno do grau médio. Dessa forma, o modelo de rede aleatória não é apropriado para representar as redes reais.

As redes sem escala, por sua vez, seguem uma distribuição denominada lei de potência, na qual nota-se que a maioria dos nós possuem poucos links em contraste com a existência de poucos nós possuindo um número elevado de ligações. O mecanismo gerador por trás do surgimento das distribuições de lei de potência, ou das redes sem escala, é denominado ligação preferencial, que determina, em resumo, que os novos nós adicionados à rede tendem a se conectar à nós com elevado grau de conexões.

Percebeu-se que, para algumas redes, quando representadas em um gráfico com escala log-log, os dados apontaram para a formação de linha reta, sugerindo uma distribuição de grau obedecendo ao seguinte modelo de distribuição:

$$p_{(k)} \sim k^{-\gamma}$$

A equação acima representa uma distribuição *power law* (lei de potência) e o expoente γ é seu expoente de grau (*degree exponente*), ou a inclinação da reta¹¹². Redes que possuem esse tipo de distribuição de graus são denominadas *scale-free networks* (redes sem escala). Em geral, redes sem escala apresentam $2 \leq \gamma \leq 3$ ¹¹³. Em resumo, em uma rede sem escala, a probabilidade de um nó se ligar a outro é diretamente proporcional ao seu grau. Assim, as redes livres de escala são dominadas por um número relativamente pequeno de nós denominados *hubs*. Uma característica fundamental deste tipo de rede é ser mais resistente à falhas acidentais, porém são mais vulneráveis à ataques coordenados, ou seja, ataques direcionados aos seus *hubs*.

Sabe-se que nas redes direcionadas cada nó é caracterizado por um k_{out} , representando o número de links que apontam do nó para outros nós, e um k_{in} , representando o número de outros nós que apontam para o nó selecionado. Deve-se, portanto, distinguir distribuições dos dois graus: a probabilidade de que um nó escolhido aleatoriamente aponte para outro k_{out} nós, ou $p k_{out}$, e a probabilidade de que um nó escolhido aleatoriamente tenha outro k_{in} nós apontando para ele,

112 Usando a forma logarítmica) fica mais evidente que o expoente do grau (γ) é a inclinação da reta dada a relação linear estabelecida.

113 Todo o tratamento estatístico das características descritas nessa sessão se encontra no anexo 1.

ou pk_{in} . Em resumo, as redes cuja distribuição de graus segue uma lei de potência (*power law*) são chamadas de redes sem escala (*scale free networks*), muitas redes reais de maior importância científica, tecnológica e sociológica tem demonstrado a propriedade sem escala (BARABÁSI, 2016). Se uma rede for direcionada, a propriedade sem escala se aplica separadamente aos graus de entrada e saída. Para estudar matematicamente as propriedades de redes sem escala, pode-se usar o formalismo discreto ou o formalismo contínuo. A propriedade sem escala é independente do formalismo usado. Além disso, quanto mais nós uma rede sem escala tem, maiores são seus hubs. Na verdade, o tamanho dos hubs cresce polinomialmente com o tamanho da rede, portanto, eles podem crescer muito em redes de escala livre.

Sumariamente, o nome sem escala captura a falta de uma escala interna, uma consequência do fato de que nós com graus muito diferentes coexistem na mesma rede. Essa característica distingue redes sem escala daquelas em que todos os nós têm exatamente o mesmo grau ($\sigma = 0$), ou de redes aleatórias, cujos graus variam em uma faixa estreita ($\sigma = \langle k \rangle^{1/2}$). Sabe-se, das propriedades estatísticas desse tipo de distribuição, que distribuições *power law* possuem variância e desvio padrão tendendo ao infinito quando o número de elementos da rede é elevado, sendo pouco informativos (BARABÁSI, 2016). Ou seja, quando escolhido um nó aleatoriamente, não se sabe qual grau esperar, traduzindo-se numa ausência de escala interna significativa, por isso “sem escala”. Consequentemente, são caracterizadas por serem *ultra small*. Essa divergência é a origem de algumas das propriedades mais intrigantes das redes sem escala, desde sua robustez até falhas aleatórias e disseminação anômala de vírus.

A propriedade *scale-free* afeta as distâncias das redes. Redes sem escala apresentam a propriedade de *small words* (ou *six degrees*), pois a presença de *hubs* reduz o comprimento da trajetória entre os nós, ao se conectarem com um grande número de nós com poucos links, criando uma pequena distância entre eles. Consequentemente, quanto mais pronunciados os hubs são, mais efetivamente eles reduzem as distâncias entre os nós, caracterizando algumas redes como *ultra small words*.

O surgimento de redes sem escala é associado a dois mecanismos: crescimento, que deriva do fato de que a maioria das redes reais não surge em sua forma final, ou seja, com a quantidade definitiva de nós todos incluídos, apresentando crescimento constante. Dito de outra forma, a rede vai se configurando na medida em que novos nós são incluídos. E a ligação preferencial, que determina que a maioria dos novos nós das redes reais prefere se conectar aos nós mais

conectados. Crescimento e fixação preferencial desempenham um papel particularmente importante na formação da distribuição de grau de uma rede, principalmente na determinação de uma distribuição de grau do tipo lei de potência. Um exemplo é uma rede de citação de artigos científicos, na medida em que novos artigos vão sendo acrescentados percebe-se a existência de alguns poucos artigos que possuem um número elevado de conexões em detrimento dos demais. Em geral, tende-se a ler os artigos mais comentados, que, geralmente, são os artigos com o maior número de citações e, por sua vez, citá-los novamente. Assim cria-se um viés no número de citações desse artigo. Quanto mais cedo um nó é adicionado a rede maior é seu grau, pois possui maiores chances de realizarem conexões futuras, o fenômeno da vantagem do primeiro a se mover.

Em resumo, nas redes reais os nós chegam um após o outro, o que altera a dinâmica da evolução de uma rede. Isso gera uma competição por links durante a qual os nós mais antigos levam vantagem sobre os mais jovens, eventualmente se transformando em hubs.

Na realidade, a taxa de crescimento de um nó não depende apenas de sua idade. Em vez disso, páginas da web, empresas ou atores têm qualidades intrínsecas que influenciam a taxa em que adquirem links. Alguns chegam tarde e, no entanto, obtêm um número extraordinário de links em um curto espaço de tempo. Outros chegam cedo, mas nunca performam bem. O fato é que as diferenças na capacidade do nó de adquirir links afetam a topologia da rede. Indo além desse cenário competitivo, outros processos, como a exclusão de nós e links ou o envelhecimento dos nós, fenômenos frequentemente observados em redes reais, também mudam a forma como as redes evoluem e alteram sua topologia. Ou seja, alguns nós, independente de quando chegaram na rede, parecem atrair mais links do que os demais. Segundo Barabási (2016) essa propriedade pode advir do grau de inovação que o nó representa, por exemplo.

A partir da metodologia de análise de redes apresentada nessa seção, o próximo capítulo se debruçará sobre as redes de patentes e artigos, descritas anteriormente, para investigar suas propriedades e características, a fim de identificar estruturas semelhantes à de uma *General Purpose Technology*.

CAPÍTULO 4: A *WORLD WIDE WEB* COMO UMA *GENERAL PURPOSE TECHNOLOGY*.

O presente capítulo se ocupará da investigação da trajetória tecnológica da WWW através da análise do padrão de formação das redes de patentes e artigos, compostas por suas respectivas citações, descritas no capítulo anterior, buscando identificar comportamentos, estruturas e componentes que se assemelhem à dinâmica das *general purpose technologies*, como apresentada e discutida no primeiro capítulo. A hipótese subjacente é a de que o conhecimento tecnológico, presente no artigo de Berners-Lee (1992), possui relevância tal que causou a explosão de numerosos progressos tecnológicos posteriores e que esse movimento poderia ser captado pela onda de patenteamento e produção científica subsequente que citaram o artigo.

Para tanto, a primeira seção se dedicará a análise descritiva dos dados compostos nas redes, como um primeiro olhar sobre o comportamento destes, para identificar algumas características e padrões na difusão do conhecimento tecnológico e científico a partir do surgimento da WWW. Em seguida, a segunda seção se debruçará sobre a análise das redes propriamente dita, a partir do arcabouço teórico discutido também no capítulo anterior, a ciência de redes. Com a inferência sobre qual tipo de topologia de rede os dados conformam e a investigação das principais características que emergem a partir dessa topologia específica, buscar-se-á encontrar elementos que corroborem o argumento de que a WWW é uma tecnologia de propósito geral, a partir da identificação destes com os discutidos na seção sobre as GPTs. E, por fim, a terceira seção se destina a uma elaboração do argumento, objetivo primeiro da presente Tese, de que a invenção da WWW conformou um sistema complexo de tecnologias com características próprias de uma GPT. Essa tarefa se desenrolará a partir da interseção entre as evidências empíricas encontradas nas duas primeiras sessões, a teoria sobre GPTs e o conhecimento disponível acerca da trajetória da tecnologia da WWW, a fim de conclusão do trabalho.

4.1. Análise descritiva dos dados: investigação preliminar.

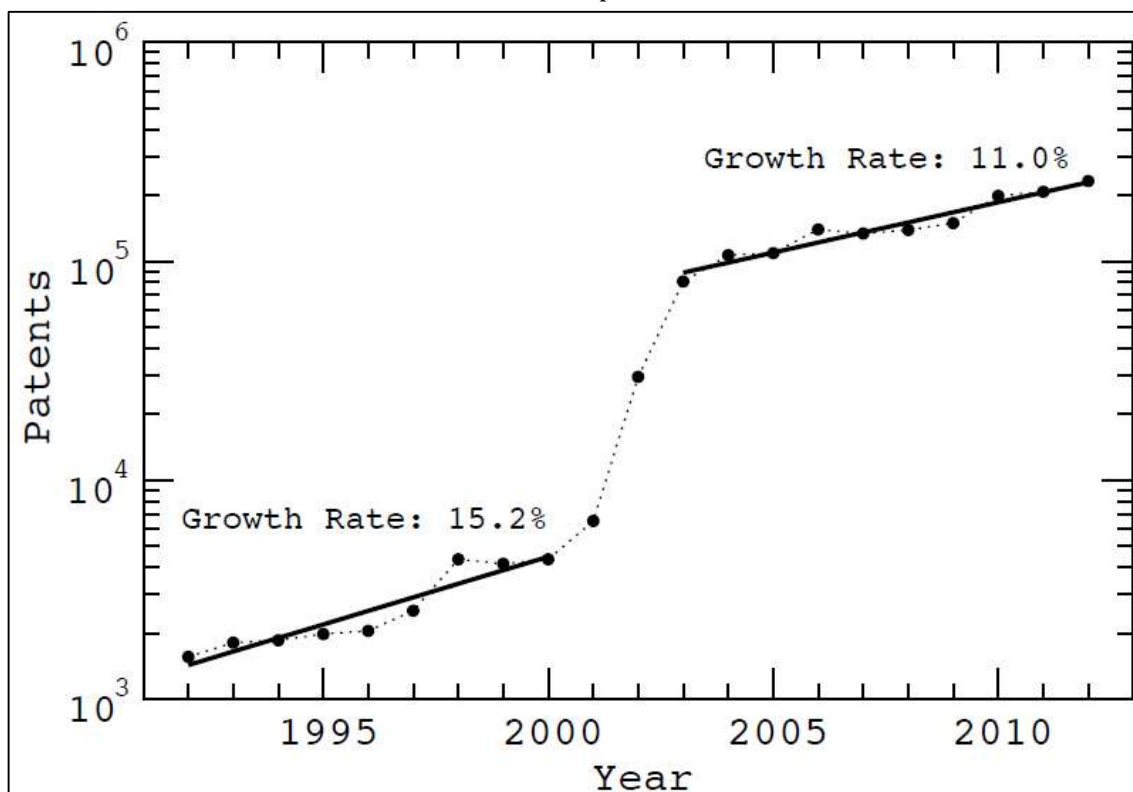
Como mencionado anteriormente, a presente seção se ocupará de uma análise descritiva dos dados que compõe as redes, ou seja, as patentes e os artigos. O objetivo é identificar algumas características na dinâmica de patenteamento e publicação de artigos, no que tange à velocidade da difusão do conhecimento e a extensão dessa difusão tanto qualitativa quanto quantitativamente. Ou seja, a partir do conhecimento inicial, expresso no artigo que propôs a WWW, quantas patentes e artigos surgiram e estão associados a ele, quais setores se

apropriaram desse conhecimento para geração subsequente de conhecimento técnico e científico, quais países lideram esses movimentos, entre outras questões relevantes.

4.1.3. Rede de patentes.

O gráfico 1, apresentado abaixo, representa a trajetória de crescimento do número de patentes concedidas dentro da rede em análise. O primeiro período de crescimento de patenteamento compreende os anos entre 1992 e 2000 onde se observa uma taxa de crescimento de 15,2%, como salientado pela reta. Percebe-se que no ano de 1992, o primeiro ano de observação da rede – lembrando que a invenção da WWW data de 1992, Berner-Lee (1992) –, conta com aproximadamente 1000 patentes. A rapidez com que o conhecimento tecnológico contido nesse artigo se espalhou pode ser um indicativo de sua característica de inovação radical ou GPT. Já nos anos 2000, o número de patentes concedidas por ano, dentro da rede, já ultrapassa a marca de 5.000 patentes.

Gráfico 1 – Crescimento do patenteamento na rede.



Fonte: elaboração própria.

A partir dos anos 2000 até, aproximadamente, o ano de 2003, ocorreu um crescimento expressivo ou exponencial do número de patentes concedidas ao ano, partindo de um número de 5.000 patentes para alcançar algo na ordem de 100 mil patentes. Mais especificamente, esse crescimento se concentra entre os anos de 2001 e 2003. Este aumento considerável do número de patentes contidas dentro da rede WWW, ocorre no mesmo período do surgimento do mecanismo de busca da Google, que popularizou o uso da internet, (Albuquerque, 2019). Antes deste mecanismo, era necessário saber o endereço exato do site que se desejava acessar, assim como o que se desejava encontrar. Com a disponibilização do mecanismo tornou-se mais rápido, intuitivo e democrático o acesso ao conteúdo online. Como mencionado no capítulo 2, o surgimento de empresas como Google, representou a início e a consolidação de um novo modelo de negócios baseado em algoritmos de identificação e discriminação do público-alvo cada vez mais elaborados. Ademais, esse período também testemunhou a crise do dot.com, com a deflagração de uma bolha financeira especulativa advinda da aposta dos investidores em oportunidades associadas às tecnologias relacionadas à WWW. Vale ressaltar que, como defendido por Perez (2002), o período inicial de difusão de uma inovação radical, ou uma GPT, em geral, apresenta um período de alta de incerteza e aumento de investimento especulativo nas novas tecnologias, resultando em crise financeira.

A partir dos anos 2003, o que se nota é uma redução da taxa de crescimento do número de patentes, com uma estabilização em torno de uma taxa, aproximada, de 11% ao ano até o ano de 2012. No entanto, apesar dessa diminuição no ritmo, o aumento de patentes continua expressivo, pois a nova taxa se dá sobre uma base inicial de 100 mil patentes. Em 2012, o último ano de observações da rede, a quantidade de patentes concedidas ao ano alcança, aproximadamente, o montante das 300 mil patentes, sinalizando uma tendência de contínuo crescimento nos próximos anos. Este fato pode indicar a importância, bem como sua prevalência ao longo do tempo, do conhecimento tecnológico representado inicialmente pelo trabalho do Bernes-Lee. Outro aspecto revelador dessa importância, é a ocorrência de patentes citando diretamente o artigo no ano de 2010, quase vinte anos após sua publicação. Além do mais, vale salientar que o gráfico do perfil de crescimento da rede WWW se assemelha bastante com o gráfico padrão de difusão de tecnologia GPT, como discutido no capítulo 1.

Com o objetivo de enxergar se houve difusão do conhecimento entre as diversas áreas tecnológicas e qual seu padrão, nos debruçamos sobre a classe tecnológica contida nos documentos de patentes. O documento de patente, além de outras informações, contém uma

classificação tecnológica da invenção reivindicada, que é atribuído pelo respectivo escritório de depósito da patente. Os registros de patentes retirados do PATSTAT, por sua vez, apresentam a classificação tecnológica internacional, o IPC, contendo classes e subclasses, o que, devido ao seu nível de detalhamento, gera um volume enorme de classificações. Para tanto, na construção dos dados, utilizou-se a conversão para os subdomínios tecnológicos OST, que correspondem às grandes áreas de conhecimento tecnológico. O conceito de subdomínios tecnológicos foi proposto pelo Observatoire des Sciences et des Techniques (OST), e permite classificar as patentes, no que tange a sua classe tecnológica, em 30 áreas de conhecimento tecnológico. A completa classificação desses subdomínios se encontra na tabela abaixo. É importante salientar que, como cada documento de patente pode apresentar mais de uma classe e subclasse tecnológica, quando da conversão para os subdomínios OST temos a possibilidade de uma mesma patente aparecer em mais de um subdomínio devido à interdisciplinaridade contida na patente.

A *Tabela 1* apresenta a distribuição das patentes ao longo dos 30 subdomínios tecnológicos OST, sendo dividida em três períodos: o primeiro compreende os anos entre 1992 e 2000; o segundo, referente ao intervalo entre 2000 e 2004; e o terceiro, os anos após 2004 até 2012. Essa divisão segue os padrões de crescimento do número de patentes apresentados no gráfico 1, representando as três fases discutidas anteriormente. Na primeira coluna da esquerda se encontram os nomes correspondentes aos 30 subdomínios tecnológicos OST, como discutido acima. Uma primeira observação relevante seria a presença de patentes em todos os subdomínios tecnológicos, desde o primeiro período observado, o que pode ser um indicativo da relevância da tecnologia inicial em basear a criação de conhecimento tecnológico nas mais diversas áreas ou do seu *spillover*, podendo indicar, em última instância, o grau de generalidade da tecnologia inicial. Uma das tendências que se observa também, é que o número de patentes, em todos os subdomínios, aumenta nos três períodos, ou seja, a quantidade de inovações que surgem, tendo como base direta, ou indireta, o artigo da formalização e divulgação da WWW, continua aumentando em todas as áreas de conhecimento ao longo dos vinte anos subsequentes à sua publicação. Ou seja, como discutido no capítulo 3, o uso de citação de patentes se justifica pela premissa da existência de uma relação tecnológica entre patentes citadas e citantes, indicando a importância tecnológica do conhecimento tecnológico citado e sua consequente criação de uma base de conhecimento para a criação de conhecimento posterior. Assim sendo, um primeiro olhar sobre a tabela 1, parece indicar uma externalidade positiva generalizada e crescente nos mais diversos setores a partir da tecnologia da WWW. Uma das características

definidoras das GPTs é sua generalidade, ou seja, possuir a capacidade de ser aplicada nos mais variados e numerosos setores.

Tabela 1 – Número de Patentes por subdomínios tecnológicos

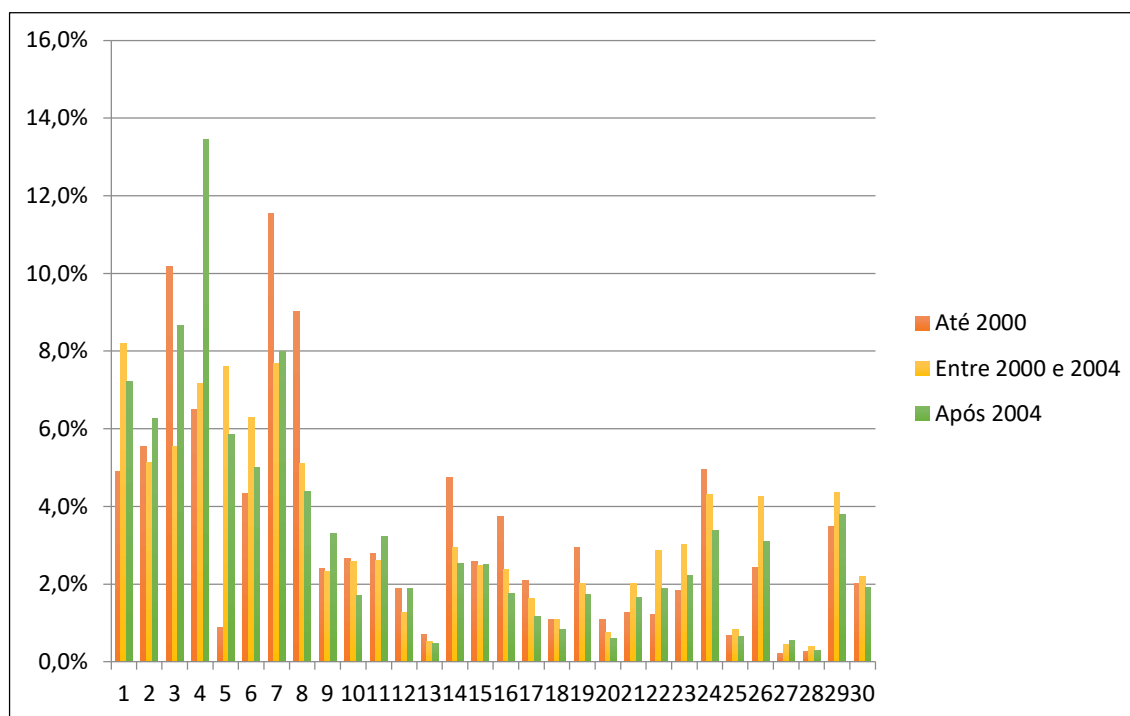
Subdomínios tecnológicos OST	Até 2000		2000 até 2004		Após 2004	
	Patentes	%	Patentes	%	Patentes	%
Componentes elétricos	1.877	4,9%	28.469	8,2%	150.553	7,2%
Audiovisual	2.116	5,5%	17.874	5,1%	130.684	6,3%
Telecomunicações	3.892	10,2%	19.243	5,5%	180.256	8,7%
Tecnologia de informação	2.488	6,5%	24.885	7,2%	279.933	13,4%
Semicondutores	334	0,9%	26.430	7,6%	121.666	5,8%
Ótica	1.658	4,3%	21.907	6,3%	104.498	5,0%
Análise, mensuração e controle	4.410	11,5%	26.734	7,7%	166.535	8,0%
Engenharia médica	3.448	9,0%	17.785	5,1%	91.195	4,4%
Produtos químicos orgânicos finos	922	2,4%	8.116	2,3%	69.027	3,3%
Química macromolecular	1.020	2,7%	9.023	2,6%	35.493	1,7%
Produtos farmacêuticos e cosméticos	1.062	2,8%	9.109	2,6%	66.941	3,2%
Biotecnologia	726	1,9%	4.386	1,3%	39.397	1,9%
Produtos agrícolas e alimentares	269	0,7%	1.784	0,5%	9.990	0,5%
Procedimentos técnicos	1.816	4,7%	10.232	2,9%	52.529	2,5%
Tecnologia de superfície e revestimento	983	2,6%	8.659	2,5%	52.424	2,5%
Processamento de materiais	1.429	3,7%	8.237	2,4%	36.470	1,8%
Materiais e metalurgia	800	2,1%	5.633	1,6%	24.227	1,2%
Técnicas térmicas	415	1,1%	3.801	1,1%	17.208	0,8%
Processamento químico básico	1.122	2,9%	7.018	2,0%	36.123	1,7%
Meio ambiente e poluição	421	1,1%	2.646	0,8%	12.585	0,6%
Máquinas-ferramentas	484	1,3%	7.023	2,0%	34.491	1,7%
Motores, bombas e turbinas	466	1,2%	9.932	2,9%	39.485	1,9%
Componentes mecânicos	705	1,8%	10.483	3,0%	46.293	2,2%
Manuseio e impressão	1.890	4,9%	15.001	4,3%	70.253	3,4%
Máquinas agrícolas e alimentares	264	0,7%	2.939	0,8%	13.872	0,7%
Transporte	931	2,4%	14.815	4,3%	64.548	3,1%
Engenharia Nuclear	81	0,2%	1.566	0,5%	11.386	0,5%
Tecnologia espacial e armas	105	0,3%	1.388	0,4%	6.073	0,3%
Bens de consumo e equipamentos	1.330	3,5%	15.169	4,4%	79.146	3,8%
Engenharia civil e construção	768	2,0%	7.661	2,2%	39.723	1,9%
Total	38.232	100%	347.948	100%	2.083.004	100%

Fonte: Elaboração própria.

O gráfico 2 apresenta os movimentos de patenteamento ao longo dos subdomínios para os três períodos delimitados, como uma proporção do total de patentes contidas na rede em cada

período. Ou seja, busca-se entender o padrão de patenteamento dentro da rede, ao longo dos subdomínios, durante o período total considerado. O eixo horizontal representa os 30 subdomínios OST, como apresentado na *Tabela 1*, e o eixo vertical as respectivas taxas percentuais.

Gráfico 2 - Evolução das patentes por subdomínios tecnológicos



Fonte: elaboração própria.

Pode-se perceber, num primeiro olhar, alguns padrões de comportamento dos movimentos de patentes na rede, como, por exemplo, o grupo composto pelos 8 primeiros subdomínios que possuem uma porcentagem significativa do total de patentes para cada período. De fato, esse grupo possui respectivamente 52,9%, 52,7% e 58,8% do total de patentes para cada período. O grupo comporta os subdomínios referentes às seguintes áreas tecnológicas: componentes elétricos, audiovisual, telecomunicações, tecnologia da informação, semicondutores, óticas, análises, medidas e controle e engenharia médica. Essa composição pode ajudar a explicar a magnitude do número de patentes pertencentes ao grupo, dado que são áreas tecnológicas, primeiro, profundamente relacionadas à chamada era da informação atual ou à quinta onda e, segundo, sendo áreas potencialmente diretamente alavancadas pela disponibilidade de uma tecnologia como a WWW, que cria, por exemplo, a necessidade de desenvolvimento de novos processadores, servidores etc., na medida que seu uso se difunde. Assim como, com potencial,

também, para aprofundar a difusão desta invenção. Ou seja, o surgimento de mecanismo de busca da Google e de computadores pessoais aprofunda a utilização da WWW em todo mundo, assim como a WWW permite o surgimento de empresas como a Google e sua nova mercadoria, bem como a sofisticação de computadores e o surgimento de novos *gadgets*.

O primeiro período, compreendendo os anos entre 1992 e 2000, tem os subdomínios relacionados à telecomunicação, análise, medição e controle e engenharia médica, com os maiores número de patentes, 10,2%, 11,5% e 9% respectivamente, sendo responsáveis por 30,7% do total de patentes no período. Apesar da clara concentração do número de patentes no grupo dos primeiros 8 subdomínios, ainda podemos perceber áreas como procedimentos técnicos e manuseio e impressão com percentuais de patentes semelhantes a alguns domínios do grupo analisado.

No segundo período, entre 2000 e 2004, percebe-se uma mudança na distribuição do número de patentes entre os subdomínios, como, por exemplo, o três com a maior porcentagem de patentes possuindo 23,5% do total. Uma diminuição do desvio padrão que pode estar relacionada ao aumento exponencial de patentes nesse período permitindo uma maior representatividade de todos os subdomínios. Além do mais, a área com maior número de patentes passa ser a de componentes elétricos, seguida pelas de análise, medição e controle e semicondutores. Esta última crescendo, aproximadamente, 750% em relação ao primeiro período. O que pode estar diretamente relacionado à explosão do uso de computadores pessoais e celulares no período. Nota-se ainda, três outros subdomínios com número expressivo de patentes neste período a saber: o de manuseio e impressão, o de transporte e de bens de consumo e equipamento.

Já o terceiro e último período, após o ano de 2004, apresenta um retorno da maior concentração de patentes nos primeiros subdomínios com as áreas da tecnologia da informação, telecomunicações e análise, medição e controle, contribuindo com 30,1% das patentes no período. No entanto, o que se nota é a liderança isolada da tecnologia da informação.

Ainda, analisando o comportamento de cada subdomínio dentro do grupo dos 8 primeiros domínios, alguns pontos podem ser levantados: o subdomínio 4, representante do subdomínio tecnológico da tecnologia da informação, que possui uma proporção significativa do número de patentes nos dois primeiros períodos – 6,5% e 7,2% respectivamente - , mas estando,

aproximadamente 44% abaixo do subdomínio com o maior número de patentes no primeiro período (análise, medição e controle), apresenta uma elevação considerável no terceiro período, possuindo, sozinho, 13,4% do total de patentes para os anos após 2004, estando quase 54% acima do subdomínio com o segundo maior número de patentes, o de telecomunicações, no período analisado. Este último, por sua vez, apresenta uma queda considerável no segundo período de quase 50%, volta a aumentar sua participação no último período. Outra dinâmica interessante é percebida no subdomínio relacionado aos semicondutores, que no primeiro período analisado possui uma parcela diminuta do total de patentes, 0,9%, aumenta sua participação na ordem de 750%, como mencionado anteriormente, e, apesar de uma ligeira queda, ainda permanece com uma taxa expressiva no último período. E, por último, vale notar que a área de engenharia médica, possui uma perda de participação no total de patenteamentos ao longo dos três períodos analisados.

Outro grupo que apresenta uma dinâmica similar entre os subdomínios é o composto por produtos químicos orgânicos finos, química macromolecular, farmacêutica e cosméticos, biotecnologia e produtos agrícola e alimentares. Este grupo se caracteriza por uma dinâmica relativamente estável da participação no total de patenteamentos ao longo dos três períodos. Apesar de, como mostra a tabela 1, apresentar um aumento do número de patentes ao longo do tempo. E por fim, nota-se o grupo que engloba os subdomínios relacionados à procedimentos técnicos, tecnologia de superfície e revestimento, processamento de material, materiais e metalurgia, técnicas térmicas, processamento químico básico e meio ambiente e poluição. Este grupo se caracteriza por uma tendência à diminuição da participação no total de patenteamento por período ao longo dos anos analisados.

Ainda vale citar os casos isolados dos subdomínios 25, 27 e 28 correspondentes às áreas de maquinário agrícola e alimentar, engenharia nuclear e de tecnologia espacial e armamentos. Tais áreas possuem uma participação reduzida no total de patenteamento nos três períodos, o que levanta a questão da representatividade da atividade inovativa dessas áreas por indicadores de patentes. Estas áreas podem configurar o caso de agentes que fizeram a opção por outro tipo de proteção intelectual, dado que a patente garante um monopólio, por prazo determinado, em troca da divulgação do conhecimento tecnológico. No entanto, a presença de patentes em todos os subdomínios parece indicar uma difusão do conhecimento desenvolvido na WWW para as demais áreas tecnológicas, não permanecendo restrito às áreas diretamente relacionadas às TICs. Ou seja, parece haver uma influência inovativa em quase todas as áreas tecnológicas. O

que permite realizar uma ponte entre esse formato de difusão da tecnologia com o previsto no modelo teórico de *general purpose technologies*, no qual a generalidade, característica fundamental da tecnologia, garante sua aplicabilidade nos mais diversos setores de atividades da economia.

Após a análise da dinâmica de patenteamento dentro a rede ao logo dos três períodos definidos, vale, ainda, algumas considerações sobre a dispersão geográfica das patentes contidas na rede e, em última instância, sobre localização da atividade tecnológica. Como mencionado anteriormente, uma das informações contidas no documento de patente diz respeito ao autor da patente, geralmente o inventor, bem como seu endereço. A prática comum, quando da utilização de dados de patentes, tem se constituído em utilizar o endereço do primeiro autor – dado que a patente pode conter mais de um – para inferir a principal localização da atividade tecnológica. Dessa forma, procedeu-se uma nova reorganização das patentes contidas na rede estudada, mas, agora, de acordo com o país citado no endereço do primeiro autor. O objetivo é tentar inferir se a invenção da WWW, promoveu ou propiciou uma difusão da atividade inovativa para além das fronteiras do país de origem e quão significativa foi essa difusão do conhecimento tecnológico. Para uniformizar a análise, classificamos as patentes de acordo com o país do primeiro autor para os três períodos utilizados anteriormente. O resultado dessa contagem de patentes por período para países, mostra que, para o primeiro período (1992 até 2000), somam-se 819 patentes distribuídas entres 27 países, o segundo período (entre 200 até 2004) contam com um total de 202.261 patentes ao longo de 102 países e, finalmente, para o último período (depois de 2004 até 2012) têm-se o número de patentes na ordem de 1.255.780 com representatividade em todos os 151 países. O que pode ser um indicativo da dispersão global do conhecimento tecnológico inicialmente gerado pela invenção da WWW.

A tabela 2 abaixo apresenta um resumo dessa dinâmica para os 20 países que obtiveram o maior número de patentes ao longo dos três períodos. Pode-se perceber algumas dinâmicas interessantes, como a manutenção de primeiro lugar dos Estados Unidos, bem como o de segundo lugar pelo Japão, ao longo do tempo. A perda de posição da Grã-Bretanha, o surgimento da Coréia e Taiwan entre os 20 países com maior número de patentes a partir do segundo período – alternando a quarta e quinta colocação, com a válida menção de que, no primeiro período, ambas as nações se faziam presente, mas com apenas 1 patente cada. Outro caso interessante é o da China, que no primeiro período não consta da lista de países com

patentes e surge, no segundo período, entre os vinte principais países, alcançando a 9ª posição no terceiro período.

Tabela 2 - Evolução das patentes por países

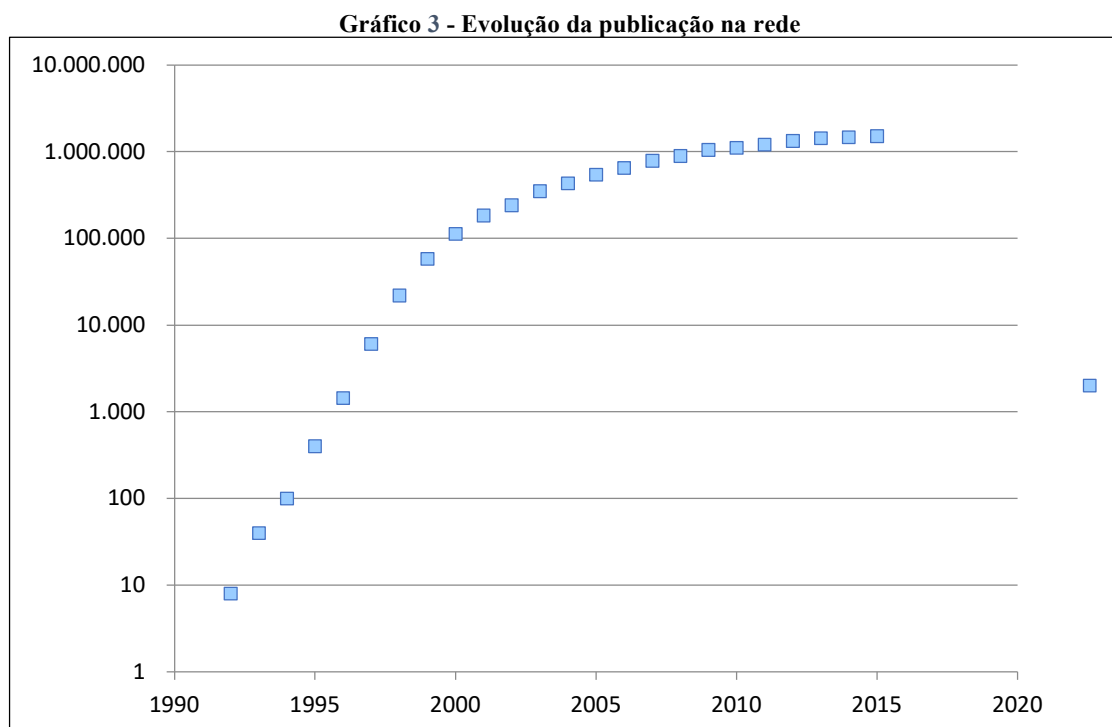
Até 2000			Entre 2000 até 2004			Depois 2004		
País	Patentes	%	País	Patentes	%	País	Patentes	%
US	399	48,72%	US	92.518	45,74%	US	589.636	46,95%
JP	90	10,99%	JP	54.207	26,80%	JP	279.727	22,28%
GB	56	6,84%	DE	13.937	6,89%	DE	75.255	5,99%
DE	51	6,23%	TW	6.395	3,16%	KR	63.187	5,03%
FR	36	4,40%	KR	6.329	3,13%	TW	47.432	3,78%
SE	27	3,30%	CA	4.428	2,19%	CA	27.954	2,23%
AU	24	2,93%	FR	4.270	2,11%	FR	26.247	2,09%
NL	24	2,93%	GB	4.169	2,06%	GB	26.062	2,08%
CA	17	2,08%	IT	2.266	1,12%	CN	14.284	1,14%
FI	16	1,95%	CH	1.709	0,84%	IT	10.999	0,88%
BE	13	1,59%	NL	1.580	0,78%	IL	10.344	0,82%
CH	13	1,59%	SE	1.327	0,66%	NL	10.325	0,82%
IT	12	1,47%	IL	1.226	0,61%	AU	9.903	0,79%
IL	9	1,10%	AU	943	0,47%	SE	9.520	0,76%
AT	6	0,73%	BE	847	0,42%	CH	9.242	0,74%
DK	5	0,61%	FI	824	0,41%	FI	6.413	0,51%
ES	4	0,49%	AT	755	0,37%	IN	4.969	0,40%
NZ	4	0,49%	CN	611	0,30%	BE	4.521	0,36%
ZA	3	0,37%	SG	595	0,29%	AT	4.256	0,34%
NO	2	0,24%	DK	469	0,23%	DK	3.632	0,29%

Fonte: elaboração própria.

A partir da observação do período de adição de patentes à rede considerada, ou seja, sua expansão temporal, bem como seu alcance nas mais diversas áreas do conhecimento e, como visto acima, na dimensão espacial, pode-se inferir, mesmo que inicialmente, que o conhecimento expresso no artigo de Berners-Lee, parece iniciar uma tendência inovativa, ou, como definido por Dosi (1982 e 1988), uma trajetória tecnológica específica que dá origem ao surgimento de numerosas tecnologias, o que é captado pela onda de patenteamento aqui analisada. Dado que a expansão temporal (sua permanência como indutor de posteriores inovações ao longo do tempo), espacial e do espectro de conhecimento (inferido a partir dos subdomínios tecnológicos) da rede sugere uma generalidade e significativa pervasividade desta tecnologia, reforça-se a hipótese de ser a WWW uma GPT típica.

4.1.4. Rede de artigos.

O gráfico 3, apresentado em seguida, representa o padrão de crescimento de publicações constantes da rede ao longo dos 23 anos considerados para sua construção. A taxa de crescimento de publicações na rede apresenta um comportamento exponencial expressivo nos primeiros 10 anos, ou seja, no período que compreende os anos 1992 e 2002, diminuindo o ritmo a partir de 2003, aproximadamente, mas mantendo um nível elevado de publicação anual.



se os anos de 1993, 2003 e 2013 para a observação da dispersão dos artigos entre as categorias de assunto. Os 40 artigos publicados no ano de 1993, pertencentes à rede, estão distribuídos ao longo de 13 categorias de assuntos da *WoS*, prioritariamente nas áreas de psicologia. Muitos desses estudos se baseavam no conhecimento de formação de redes não hierárquicas e criação de sistemas de armazenamento de informação presentes no artigo inicial de Berners-Lee. Como descrito no capítulo 2, a ideia base do sistema de informação proposto por Berners-Lee era inspirado pela forma de funcionamento da mente humana, que liga informações de forma não hierárquica, assim, é natural que essas áreas de estudo relacionadas ao funcionamento cognitivo humano estejam presentes na rede desde seus primórdios.

Tabela 3 - Evolução da publicação por área do conhecimento.

1993		2003		2013	
Área do conhecimento	Artigos	Área do conhecimento	Artigos	Área do conhecimento	Artigos
Psicologia	13	Bioquímica e biologia molecular	41.746	Engenharia elétrica e eletrônica	53.658
Psicologia, experimental	6	Neurociências	13.701	Bioquímica e biologia molecular	50.382
Métodos de pesquisa bioquímicos	5	Oncologia	13.578	Ciências multidisciplinares	48.356
Psicologia, desenvolvimento	4	Química, multidisciplinar	12.924	Química, multidisciplinar	48.252
Psicologia, multidisciplinar	3	Imunologia	10.647	Química, física	39.216
Negócios	2	Química, física	9.551	Oncologia	32.527
Neurociências	1	Neurologia clínica	8.913	Ciências dos materiais, multidisciplinar	26.566
Psicologia, biológica	1	Farmacologia e farmácia	8.874	Neurologia clínica	24.934
Ciências comportamentais	1	Endocrinologia e metabolismo	7.852	Sistemas cardíaco e cardiovasculares	22.936
Bioquímica e biologia molecular	1	Medicina, geral e interna	7.801	Medicina, geral e interna	22.480
Economia	1	Biologia celular	7.726	Ciência da computação, Inteligência Artificial	22.021
Problemas Sociais	1	Biotecnologia e microbiologia aplicada	7.590	Ciências ambientais	21.528
Ciências veterinárias	1	Métodos de pesquisa bioquímicos	7.314	Astronomia e astrofísica	17.256

	Sistemas cardíaco e cardiovasculares	6.915	Física, aplicada	17.211
	Química, orgânica	6.094	Sistemas de automação e controle	16.977

Fonte: elaboração própria.

O ano de 2003 apresenta uma mudança significativa tanto quantitativa quanto qualitativa, pois o número de publicações sobe para 351.128 artigos, variando ao longo de 241 categorias, onde nota-se a prevalência de áreas como a bioquímica e biologia celular, neurociências, química, oncologia. Já em 2013, observa-se um quantitativo de 1.430.894 artigos divididos em 251 categorias. Surgem aqui novas áreas expressivas como engenharia elétrica e eletrônica, ciências multidisciplinares e as ciências da computação e inteligência artificial, além da permanência de diversas áreas da medicina, química, biologia e bioquímica. Observar esse padrão de publicação por áreas do conhecimento pode nos dar indícios da potencial capacidade de indução de criação de conhecimento nas demais áreas científicas bem como sua transformação ao longo do tempo. Ademais, vale salientar que os diversos conhecimentos na base da atividade inovativa possuem formas variadas de divulgação e proteção, o que pode explicar as diferenças de áreas de conhecimento entre as redes de patentes e artigos. Ainda, é importante lembrar que, como discutido no primeiro capítulo do presente trabalho, existe uma boa parcela do conhecimento empregado no processo inovativo que não pode ser captado por artigos e patentes, dada sua característica informal e tácita. A tabela 3 apresenta um ranking das 15 categorias de assuntos que possuem a maior quantidade de artigos dentro da rede, nos três anos de análise.

É importante salientar que o artigo produzido por Berners-Lee continha uma base de conhecimento sobre redes neurais e outros princípios da neurociência, justificando a presença de áreas como psicologia e neurociências como detentoras do maior número de artigos no primeiro período. Mas é clara a mudança de composição do grupo das 15 áreas com maior número de publicação, dentro da rede, ao longo do período. Cabe ainda ressaltar que cada setor de atividade inovativa apresenta padrões distintos de divulgação de conhecimento produzido de acordo com suas especificidades. Ademais, pode-se, ainda, analisar a distribuição dos artigos constituintes da rede ao longo dos variados países. Apenas 8 países aparecem na distribuição dos artigos em 1993, em 2003 o número cresce para 70, alcançando todos os países em 2013. A tabela 4 apresenta, para os mesmos anos da tabela 3, um ranking dos 15 países com o maior número de publicação dentro da rede. A distribuição regional dos artigos apresenta muitas semelhanças com a de patentes.

Assim como discutido a partir da rede de patentes anteriormente, a observação da expansão temporal, do espectro de áreas do conhecimento e da dimensão espacial das publicações científicas, corrobora a hipótese de que o conhecimento expresso no artigo de Berners-Lee parece iniciar uma tendência na produção de conhecimento, tanto técnico como científico, que alcança praticamente todos os domínios tecnológicos e científicos, que se perpetua por um período considerável de tempo e que se espalha geograficamente. Dado que a expansão temporal, espacial e do espectro de conhecimento da rede, sugere uma generalidade e significativa pervasividade deste corpo de conhecimento, pode-se inferir a presença de uma GPT, a partir dessas análises preliminares dos dados.

Tabela 4 - Evolução da publicação por países.

1993		2003		2013	
País	Artigos	País	Artigos	País	Artigos
EUA	11	EUA	118.507	EUA	296.950
Suíça	5	Japão	27.147	China	204.104
Inglaterra	1	Alemanha	24.929	Alemanha	69.401
Austrália	1	Inglaterra	21.634	Japão	66.583
França	1	França	16.249	Inglaterra	60.100
Finlândia	1	Itália	13.509	Itália	48.421
Alemanha	1	Canadá	12.990	Índia	47.459
Canadá	1	China	10.568	França	46.708
		Espanha	8.804	Canadá	43.426
		Austrália	7.904	Coréia do Sul	40.658
		Holanda	7.501	Espanha	40.085
		Suécia	5.638	Austrália	38.237
		Coréia do Sul	5.110	Brasil	32.151
		Suíça	4.941	Taiwan	24.374
		Índia	4.264	Holanda	23.966

Fonte: elaboração própria.

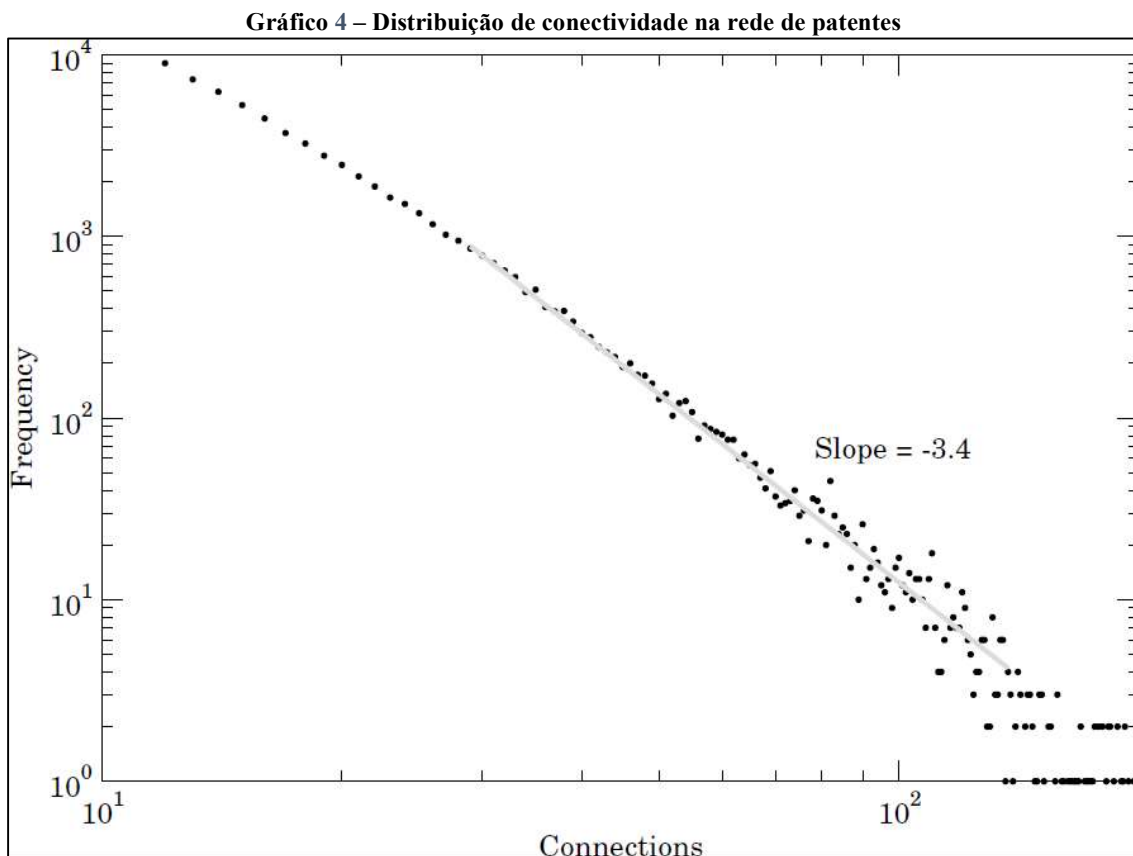
4.2. Análise das redes: *Science Networks*.

As próximas duas subseções se dedicarão à análise das redes, propriamente ditas, a partir dos ferramentais metodológicos fornecidos pela ciência de redes (BARABÁSI, 2016; NEWMAN, 2018; MENCZER ET AL., 2020; LATORA ET AL., 2017; CALDARELLI E CHESSA, 2016), discutidos no capítulo anterior. Objetiva-se analisar a topologia assumida pelas redes, que por sua vez representam os sistemas complexos tecnológico e científico, conformados com o surgimento da WWW. A partir da definição da topologia das redes e das características que tais

estruturas acarretam, será possível traçar um paralelo entre o comportamento da WWW e a dinâmica própria de uma GPT. Para tanto, a primeira subseção se ocupará da análise da rede de patentes e a segunda da rede de artigos.

4.2.3. Rede de patentes da World Wide Web.

A rede de patentes e suas respectivas citações aqui analisada constitui-se de nós representando as patentes e dos links representando as citações entre elas. A distribuição de links (distribuição de grau ou *degree distribution*) é representada no gráfico 4 abaixo, no qual a distribuição de graus é mostrada em um gráfico log-log, plotando-se a frequência em função do número de conexões.



Fonte: elaboração própria.

A rede em questão é uma rede direcionada e o gráfico 4, acima, representa a distribuição dos *in-degrees*, ou seja, a distribuição de frequência dos links que cada nó recebeu, ou ainda, das citações que cada patente recebeu de patentes subsequentes. Isso é um resultado direto da metodologia utilizada na construção da rede, onde cada novo nível é composto por patentes que

citaram as patentes subjacentes, apresentada no capítulo anterior. A rede direcionada analisada possui $N = 1.569.977$ patentes como nós e $L = 6.320.366$ citações como links.

Apesar do grau médio dos nós ser igual a 4,5, ou seja, cada patente recebeu, em média 4.5 citações de patentes subsequentes, a distribuição de graus p_k , mostrada no gráfico 4, indica que a grande maioria dos nós possui apenas alguns links. Esses numerosos nós com poucos links coexistem com alguns poucos nós ou hubs altamente conectados, o maior com até 505 links. Essas grandes diferenças em graus de nós são uma consequência da propriedade de escala livre (*free scale*) da rede. Como foi visto no capítulo anterior, a forma da distribuição de graus determina uma ampla gama de propriedades de rede, desde a robustez da rede até a propagação de eventos específicos.

Ainda de acordo com o exposto no capítulo 3, nas redes que, em geral, quando representadas em um gráfico com escala log-log, os dados apontarem para a formação de linha reta, podem sugerir uma distribuição de grau obedecendo ao modelo de distribuição *power law*, representada pela equação: $p_k \sim k^{-\gamma}$. Sendo o expoente γ o seu expoente de grau (*degree exponente*), ou a inclinação da reta. Redes que possuem esse tipo de distribuição de graus são denominadas *scale-free networks* (redes sem escala). E, geralmente, redes sem escala apresentam o valor do expoente de grau variando entre $2 \leq \gamma \leq 3$.

Portanto, após o cálculo da inclinação da reta presente no gráfico 4, que representa a distribuição de grau da rede em questão, obtém-se que o valor desta inclinação é $\gamma = 3,4$. Ou seja, a rede em questão apresenta uma distribuição do tipo lei de potência (*power law*), com o expoente de grau igual a $\gamma = 3,4$. É importante salientar que, como argumenta Barabási (2016), redes complexas que representam sistemas complexos que ainda estejam em formação, ou seja, redes nas quais novos nós e conexões ainda podem ser adicionados, podem apresentar um expoente de grau fora do intervalo determinado pelo modelo, sendo diferente para redes já consolidadas. O importante é notar se a rede possui uma grande quantidade de nós com um número pequeno de links e outros poucos nós contando com um número elevado de conexões, ou *hubs*, que determinam a dinâmica e a topologia da rede.

Como foi visto, a propriedade *scale-free* afeta as distâncias das redes dado que apresentam a propriedade de *small worlds* (ou *six degrees*), que é uma consequência direta da presença de

hubs que reduzem o comprimento da trajetória entre os nós ao se conectarem com um grande número de nós com poucos links, criando uma pequena distância entre eles. Consequentemente, quanto mais pronunciados os hubs são, mais efetivamente eles reduzem as distâncias entre os nós, caracterizando algumas redes como *ultra small words*. Essa propriedade parece indicar maior conectividade na rede, ou seja, a potencial presença de uma tecnologia de alta pervasividade. Ou seja, a propriedade *small words* garante que se pode percorrer o caminho entre quaisquer dois nós da rede, deslocando-se ao longo de poucos links, o que garante, quando se tratando de uma rede de tecnologia, por exemplo, uma rápida e eficiente difusão do conhecimento, tanto tecnológico quanto científico, como ficará evidente adiante. Ademais, a presença dos *hubs* dá a rede um outro aspecto importante que trata da sua robustez, como discutido por Barabási (2016). A robustez, característica das redes sem escala, define que a rede é menos vulnerável a quaisquer retiradas aleatórias de nós, tornando-a uma rede de difícil desmantelamento, isto é, a difusão de conhecimento tende a permanecer por um período longo de tempo. A tabela 5 apresenta os maiores hubs da rede em questão, que se definem pelas patentes que receberam o maior número de citações:

Tabela 5 - Patentes com os maiores números de citações.

Citações	Ano	Titular	País	OST
505	2003	SanDisk Corporation	US	4,5
423	2004	Tubarc Technologies, LLC	US	24
338	2003	3M Innovative Properties Company	US	24,4
309	2003	Aisin AW Co., Ltd.	JP	23
305	2003	Saba Software, Inc.	US	4
271	2001	Veridian Information Solutions	US	2,3,4,7
256	2003	Fujitsu Limited	JP	3,4,7
251	2003	Qualcomm Incorporated	US	3
248	2002	Lucent Technologies, Inc.	US	3,4
243	2004	Silverbrook Research Pty Ltd	AU	2,24,29,4,7
235	2003	Hewlett-Packard Development Company, L.P.	US	3,4
234	2005	SanDisk Corporation	US	4
229	2002	Matrix Semiconductor, Inc.	US	4,5
229	2003	Kabushiki Kaisha Toshiba	JP	4,5
223	2003	Intel Corporation	US	5
221	2002	Macronix International Co., Ltd.	TW	4
219	2003	Hoffberg; Steven	US	4,7
216	2003	Agfa-Gevaert	BE	24
214	2003	Silverbrook Research Pty Ltd	AU	24
213	2003	Ovonyx, Inc.	US	4,5
212	2005	Tangis Corporation	US	4
211	2004	Silverbrook Research Pty Ltd.	AU	2,24
211	2005	Tubarc Technologies, LLC	US	24

210	2003	ASM Microchemistry Oy	FI	15,5
210	2003	Flarion Technologies, Inc.	US	3,4
206	2006	Friendster, Inc.	US	4
205	2004	Fujitsu Limited	JP	2
204	2003	Nikon Corporation	JP	6
201	2002	Crossroads Systems, Inc.	US	3,4
201	2004	Dainippon Screen Mfg. Co., Ltd.	JP	24,6
201	2006	Tubarc Technologies, LLC	US	24,30
200	2004	Nikon Corporation	JP	6

Fonte: elaboração própria

A tabela 5 apresenta as patentes com os maiores graus, ou número de citações recebidas, com as informações dos respectivos números de citações, nomes dos titulares, países dos titulares e os subdomínios tecnológicos aos quais pertencem. Nota-se, como esperado, uma prevalência de patentes relacionadas aos subdomínios de tecnologias de informação, telecomunicações, semicondutores, ótica, manipulação e impressão. Além de uma concentração temporal no início dos anos 2000, momento a partir do qual o tamanho da rede apresenta um crescimento significativo, como visto no gráfico 1 anteriormente. Ademais, observa-se um domínio dos EUA e Japão em sediar as empresas titulares das patentes. A título de maior detalhamento, a tabela 6 revela o tipo de tecnologia representada pelos oito maiores hubs:

Tabela 6 - Tecnologia dos maiores hubs.

Citações	Título da patente
505	Técnicas operacionais para reduzir os efeitos do acoplamento entre elementos de armazenamento de uma memória não volátil operada em vários estados de dados.
423	Condução de fluidos utilizando um sifão insaturado reversível com ação de porosidade tubarc.
338	Sistema e método de rastreamento e gerenciamento de objetos usando etiquetas de identificação por radiofrequência.
309	Laminado do dispositivo de exibição de imagem e método de fabricação do dispositivo de exibição de imagem.
305	Método e aparelho para gerenciar troca de dados entre sistemas em uma rede.
271	Sistema de controle de acesso e distribuição de propriedade digital.
256	Dispositivo de exibição de imagem.
251	Método e aparelho para transmissão de dados por pacotes de alta taxa.
248	Fornecendo transparência de recursos entre uma pluralidade de PBXs.

Fonte: elaboração própria.

Percebe-se a presença de tecnologias referentes aos novos produtos diretamente relacionados aos setores de telecomunicações, como às relacionadas aos dispositivos de exibição de imagem; de impressão, como a tecnologia de condução de fluidos para aplicação em processos de

impressão; tecnologias de informação, como as relacionadas à capacidade de armazenamento e transporte de dados etc.

Sabe-se que o surgimento de redes sem escala é associado a dois mecanismos: o crescimento e a ligação preferencial. O primeiro deles deriva do fato de que a maioria das redes reais não surge em sua forma final, ou seja, com a quantidade definitiva de nós todos incluídos, apresentando crescimento constante. Ou seja, a rede vai se configurando na medida em que novos nós são incluídos. Já a ligação preferencial determina que a maioria dos novos nós das redes reais prefere se conectar aos nós mais conectados. Esse fato pode ser um indicativo de patentes com elevada importância tecnológica, ditando os caminhos da inovação a partir do seu surgimento, configurando-se clusters de tecnologias fundamentais para o desenvolvimento e suporte da WWW, determinando paradigmas e trajetórias tecnológicas.

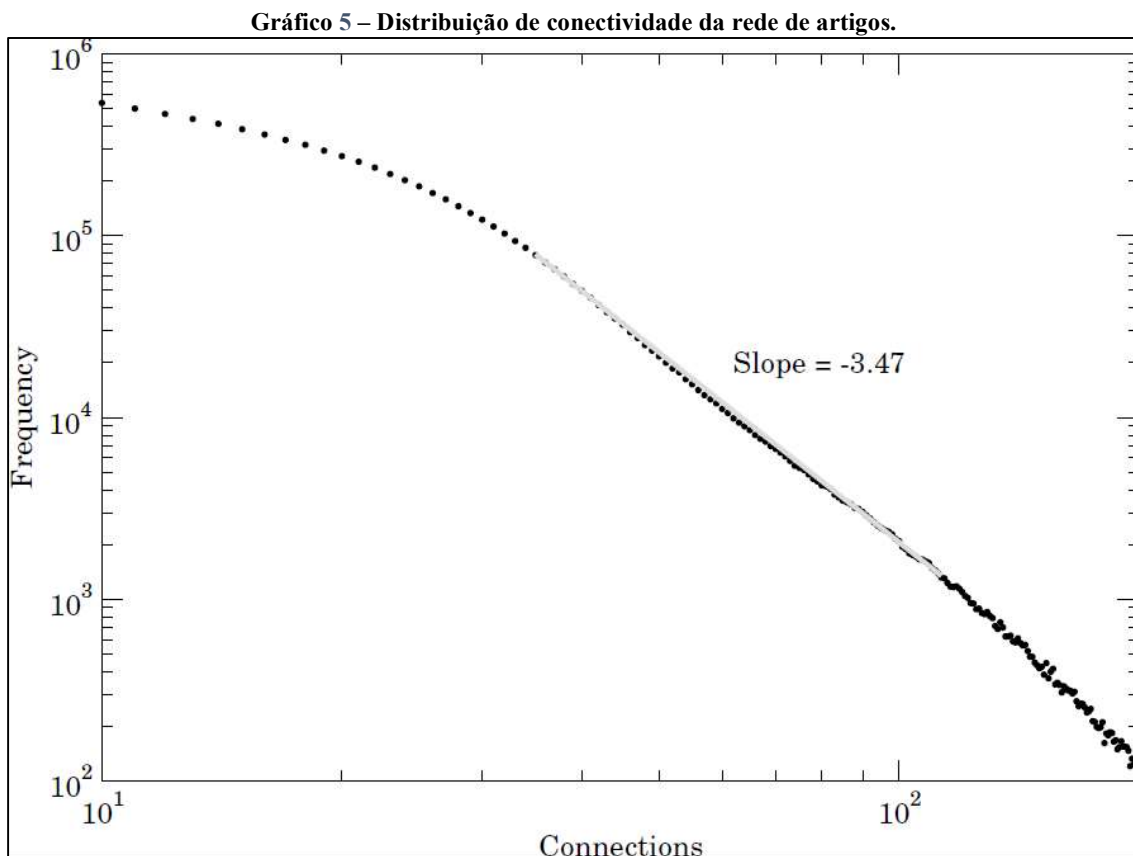
Crescimento e fixação preferencial desempenham um papel particularmente importante na formação da distribuição de grau de uma rede, principalmente na determinação de uma distribuição de grau do tipo lei de potência. Um exemplo é uma rede de citação de artigos científicos, na medida em que novos artigos vão sendo acrescentados percebe-se a existência de alguns poucos artigos que possuem um número elevado de conexões em detrimento dos demais. Em geral, tende-se a ler os artigos mais comentados, que, geralmente, são os artigos com o maior número de citações e, por sua vez, citá-los novamente. Assim cria-se um viés no número de citações desse artigo. Quanto mais cedo um nó é adicionado a rede maior é seu grau, pois possui maiores chances de realizarem conexões futuras, o fenômeno da vantagem do primeiro a se mover. Assim, nas redes reais os nós chegam um após o outro, o que altera a dinâmica da evolução de uma rede. Isso gera uma competição por links durante a qual os nós mais antigos levam vantagem sobre os mais jovens, eventualmente se transformando em hubs.

Na realidade, a taxa de crescimento de um nó não depende apenas de sua idade. Observa-se que, em geral, páginas da web, empresas ou atores têm qualidades intrínsecas que influenciam a taxa na qual adquirem links. Alguns chegam tarde e, no entanto, obtêm um número extraordinário de links em um curto espaço de tempo. Outros chegam cedo, mas nunca performam bem. O fato é que as diferenças na capacidade do nó de adquirir links afetam a topologia da rede. Indo além desse cenário competitivo, outros processos, como a exclusão de nós e links ou o envelhecimento dos nós, fenômenos frequentemente observados em redes reais, também mudam a forma como as redes evoluem e alteram sua topologia. Ou seja, alguns nós,

independente de quando chegaram na rede, parecem atrair mais links do que os demais. Segundo Barabási (2016) essa propriedade pode advir do grau de inovação que o nó representa, por exemplo, sendo uma tecnologia capaz de criar uma onda de tecnologias subsequentes com certa dependência tecnológica.

4.2.4. Rede de artigos da WWW.

A rede de artigos é constituída por $N = 13.421.547$ artigos representados pelos nós e $L = 109.288.513$ citações definindo seus links. De forma análoga ao realizado na análise da rede de patentes, plota-se a distribuição de frequência dos graus para a rede de artigos, também em função logarítmica, resultando no gráfico 5, apresentado abaixo. Nota-se a presença de muitos nós com menor número de conexões em contrapartida a poucos nós com quantidade expressiva de links. Ou seja, muitos artigos recebendo um número menor de citações em artigos subsequentes e alguns poucos artigos recebendo muitas citações. Indicando a presença de uma distribuição do tipo lei de potência, a ser confirmada pelo cálculo do expoente de grau da reta definida pela distribuição, ou, simplesmente, sua inclinação.



Fonte: elaboração própria.

O coeficiente de grau apresenta o valor de $\gamma=3,47$, muito próximo ao valor da rede de patentes, e, também, ligeiramente fora do intervalo definido pelo modelo, mas, como mencionado anteriormente, esse é um comportamento esperado para redes que ainda se encontram em formação. Dessa forma, a rede se configura como sendo do tipo sem escala (*scale free*), possuindo todas as características discutidas na sessão anterior. Vale salientar que redes compostas por publicações científicas e suas citações, em geral, apresentam esse tipo de distribuição por serem redes em formação e por apresentarem nós com propriedade de fixação preferencial, derivada diretamente da dinâmica científica, como discutido no capítulo 3. A forma de validação científica, em geral, é ser o primeiro a publicar uma grande descoberta, assim cria-se reputação e uma base sobre a qual as futuras pesquisas serão feitas. Ainda, de acordo com Dosi (1988), um paradigma tecnológico define as perguntas, bem como a forma como tais perguntas devem ser respondidas, o que naturalmente cria um viés, ou preferência, em torno de alguns tipos de conhecimento. Abaixo, têm-se os maiores hubs da rede de artigos, com as respectivas informações do número de citações recebidas, ano de publicação, instituição de origem da pesquisa, país e área de conhecimento:

Tabela 7 - Artigos com maiores números de citações recebidas.

Citações	Ano	Instituição	País	Área do conhecimento
201	2005	MIT	USA	Bioquímica e Biologia Celular
170	1999	UNIV SHEFFIELD	UK	Ciência da Computação, Ciência da informação e Ciência da Biblioteca
169	2007	UNIV AUTONOMA MADRID	ES	Oncologia
168	2008	UNIV CHICAGO	USA	Ciências Multidisciplinares
159	2013	TRINITY COLL DUBLIN	IE	Ciências Multidisciplinares
153	2009	UNIV LIBRE BRUXELLES	BE	Endocrinologia e Metabolismo
149	2010	MEM SLOAN KETTERING CANC CTR	USA	Ciências Multidisciplinares
143	2004	HARVARD UNIV	USA	Ciências Multidisciplinares
143	2010	UNIV MASSACHUSETTS	USA	Ciências Multidisciplinares
140	2003	UNIV N CAROLINA	USA	Ecologia
139	2007	ARISTOTLE UNIV THESSALONIKI	EL	Ciência e tecnologia verde e sustentável
138	2008	NATL PUBL HLTH INST	FI	Endocrinologia e Metabolismo
137	2004	HADASSAH HEBREW UNIV MED CTR	IL	Ciências Multidisciplinares
137	2008	HANNOVER MED SCH	DE	Bioquímica e Biologia Celular

136	2000	NETHERLANDS INST ECOL	NL	Limnologia
136	2005	AUSTRALIAN NATL UNIV	AU	Astronomia e Astrofísica
136	2011	UNIV FLORIDA	USA	Ecologia
133	2005	AUSTRALIAN NATL UNIV	AU	Ciências Multidisciplinares
133	2009	PENINSULA MED SCH	UK	Endocrinologia e Metabolismo
133	2009	CTR GENOM REGULAT CRG	ES	Métodos de pesquisa bioquímicos
131	2013	MCGILL UNIV	CA	Biologia Celular
131	2013	UNIV PENN	USA	Oncologia
130	2004	TOHOKU UNIV	JP	Física, Matéria condensada
130	2008	UNIV CAMBRIDGE	UK	Medicina, Geral e Interno
128	2008	UNIV GUELPH	CA	Ciências Multidisciplinares
128	2011	STANFORD UNIV	USA	Ciências Multidisciplinares

Fonte: elaboração própria.

A tabela 8 apresenta os títulos das oito publicações que mais receberam citações na rede, ou seja, dos maiores *hubs*. Nota-se uma concentração em áreas biológicas no geral e ciências da computação e informação. Nas áreas de ciências da computação e informação, a presença de hubs é esperada e natural, dada a natureza da tecnologia da WWW. No entanto, na áreas biológicas, a princípio, esse resultado pode não ser tão óbvio, até que se saiba que Berners-Lee se inspirou nos conceitos de redes neurais, ou redes de neurônios, para criar seu conceito de redes de informações. Assim, a presença de hubs nessas áreas parece ser tanto um desdobramento desta inspiração, quanto um resultado de como o conhecimento criado com a WWW se difundiu para outras áreas e criou um mecanismo de retroalimentação de criação e sofisticação do conhecimento na rede, uma das característica marcante de GPTs.

Tabela 8 - Títulos das publicações com o maior número de citações recebidas.

Citações	Título da publicação
201	Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets
170	Models in information behaviour research
169	Snail, ZEB and bHLH factors in tumour progression: an alliance against the epithelial phenotype?
168	Innate immunity and intestinal microbiota in the development of Type 1 diabetes
159	Metabolism of inflammation limited by AMPK and pseudo-starvation
153	The role of inflammation in insulinitis and beta-cell loss in type 1 diabetes
149	Skp2 targeting suppresses tumorigenesis by Arf-p53-independent cellular senescence
143	Structural basis for removal of adenine mispaired with 8-oxoguanine by MutY adenine DNA glycosylase

Fonte: Elaboração própria.

4.3. A *World Wide Web* como uma *General Purpose Technology*: uma proposta de interpretação.

É difícil exagerar a importância da World Wide Web em nossa vida diária. Da mesma forma, não podemos exagerar o papel que a WWW desempenhou no desenvolvimento da teoria de rede: facilitou a descoberta de várias características fundamentais da rede e tornou-se um teste padrão para a maioria das medidas de rede (BARABÁSI, 2016, p. 3, cap.4).¹¹⁵

A partir da investigação do processo de evolução e difusão do conhecimento tecnológico e científico derivado da invenção da WWW, representado respectivamente pelas redes de patentes e artigos e suas citações que, por sua vez, foram analisadas sob o arcabouço metodológico da ciência de redes, busca-se identificar características e propriedades nessas dinâmicas que se assemelhem ao comportamento próprio de uma GPT. No primeiro capítulo a tecnologia foi definida como sendo um conjunto de conhecimentos, tanto técnico como tácito, científico ou tecnológico (DOSI, 1982; ARTHUR, 2009), justificando a investigação por meio das redes de difusão de conhecimento científico e tecnológico. Arthur (2009) define que o processo de evolução da tecnologia determina que toda tecnologia atualmente disponível possui uma linha de dependência de tecnologias anteriores, ao definir que tecnologias são sempre a combinação de tecnologias já disponíveis (os *building blocks*). Ou seja, seria possível traçar uma linha genealógica para qualquer tecnologia, o que embasa o uso de redes para a investigação da evolução da tecnologia em questão, a WWW. Ademais, Dosi (1982 e 1988) define que o processo inovativo é acumulativo.

Nas teorias acerca de GPTs define-se que esta é uma tecnologia de aspecto geral e alta capacidade de pervasividade. Sendo uma de suas principais características a capacidade de ser aplicada, ou absorvida, nos mais diferentes setores da atividade econômica e científica, para além dos setores mais óbvios de aplicação direta, que no caso da WWW seriam os setores de computação e redes de computadores. Como foi visto, em ambas as redes, existem patentes e artigos em todos os subdomínios tecnológicos e áreas científicas, indicando a generalidade da tecnologia e sua capacidade de transbordamento para as mais diferentes áreas. O uso da ciência, ou teoria, de redes possui a vantagem de captar quem são seus componentes e os links entre eles, permitindo identificar quais os setores tecnológicos e científicos que fazem parte do

115 Tradução do original em inglês.

processo de difusão e evolução da tecnologia da WWW. Em especial, no caso da propriedade *free-scale* das redes utilizadas, sabe-se da presença de *hubs* que desempenham um papel-chave na determinação do comportamento da rede, ou seja, existem domínios tecnológicos e áreas científicas que desempenharam um papel-chave na difusão do conhecimento incorporado na tecnologia da WWW, setores sem os quais a rede não escalaria, rede que representa o sistema complexo de tecnologia inaugurado pela invenção da WWW. E, por sua vez, a existência desses setores, tecnologias e conhecimento científico, depende diretamente da existência da WWW.

Esta propriedade remete diretamente às características de identificação de uma GPT de Bekar, Carlaw e Lipsey (2018), discutidas no capítulo 1:

1. Complementaridades com um cluster de tecnologias que o definem e dão suporte: como observado pela presença dos maiores *hubs* nos subdomínios relacionados a telecomunicações, tecnologias de informação, semicondutores, ótica. São ao mesmo tempo tecnologias que definem e dão suporte a WWW, através das tecnologias disponíveis em computadores (softwares e hardwares), a internet, linguagem de programação computacional, entre outras.
2. Complementaridades com um cluster de tecnologias que ele habilita: a WWW permitiu inúmeras novas invenções e inovações, nos mais variados e numerosos setores da economia, muitas das quais tecnicamente impossíveis ou economicamente inviáveis sem a GPT e que não são identificáveis como parte da GPT (ou seja, com o cluster de tecnologias que define e dá suporte a GPT): como foi o caso do surgimento de empresas como a Google e seu novo modelo de negócio e mercadoria, a publicidade; o *ecommerce* que evoluiu para as grandes plataformas de negócios como a Amazon; redes sociais inaugurando o que se denomina como a economia da atenção; grandes plataformas de economia compartilhada como a Uber; bancos digitais e, por fim, a disponibilidade de Inteligência Artificial, Internet das coisas e Blockchains, dado o volume de dados gerado e armazenado via a conexão generalizada permitida pela WWW. Ademais, permitiu o avanço de outros setores como a neurociência e outros ramos da medicina, ao acelerar a melhora no processamento computacional e armazenamento de dados.
3. Complementaridades com um cluster de tecnologias que normalmente incluem aquelas que são social, política e economicamente transformadoras: na medida em que a GPT expande o alcance e a variedade de suas aplicações, as tecnologias envolvidas têm impactos significativos na estrutura econômica e, muitas vezes, nas estruturas políticas e sociais, exigindo mudanças ou abrindo oportunidades para inovações nas tecnologias

que são ou serão incorporados nessas estruturas. Como visto acima, tecnologias como redes sociais, plataformas online de comércio, serviços e compartilhamento (Uber, Airbnb, streaming como Netflix, Amazon Prime e Spotify, entre outros), tem alterado nossa forma de interação social, nosso comportamento de consumo, nossas relações com o trabalho, bem como demandando o surgimento de novas leis e instituições para lidar com tais alterações. Como, por exemplo, leis sobre acesso e divulgação de dados e informações dos usuários da rede, qual o uso que se pode fazer dos milhares de dados que são gerados diariamente na rede por parte das empresas; discussões sobre a precarização das relações de trabalho implícitas nas plataformas de prestação de serviço – denominada usualmente com a “uberização” do trabalho – dando lugar a novas legislações etc.

4. Não há substitutos próximos: as GPTs criam complementaridades com muitas ou a maioria de suas aplicações que são: (i) únicos -nenhuma outra combinação de tecnologias pode produzir a aplicação - e (ii) Leontief - sem ele, todo o sistema não funcionaria. Isso fica evidente a partir de como a WWW foi criada usando tecnologias já existentes, mas como resultado de uma combinação única criando todo um novo potencial universal de aplicações, antes impossível.
5. Tenha uma ampla variedade de aplicações: a ampla gama de complementaridades da GPT (pontos 1 a 3) resulta em GPTs com vários usos ou um único uso genérico que possui muitas aplicações econômicas em grande parte ou em toda a economia.
6. Inicialmente bruto, mas evoluindo em complexidade: as tecnologias que evoluem para se tornarem GPTs geralmente começam de forma bruta e incompleta, geralmente com uma gama restrita de usos. Como discutido no capítulo 2, os primeiros navegadores eram muito rudimentares e não possuíam capacidade de edição, apenas com as futuras colaborações e complementariedades e o surgimento de empreendimentos no ramo, a WWW pode se desenvolver e difundir se tornando no que conhecemos hoje.

Ademais a propriedade de *small-words* garante um caminho curto entre qualquer par de nós da rede, o que é reforçado pela presença da propriedade *free-scale*. O que pode indicar, novamente, o alto grau de generalidade da tecnologia da WWW. A presença dos *hubs*, indicando que novos nós preferem se conectar aos nós mais conectados, indicam que a característica definida por Dosi (1988) de que a inovação tende a ser localmente localizada, apontando para a existência de um paradigma tecnológico que define a heurística da atividade inovativa com suas respectivas trajetórias, está presente no padrão de difusão da WWW. Ou seja, os indivíduos

tendem a inovar dentro de áreas nos quais possuem expertise. Ou, indivíduos que inovam tendem a continuar inovando. Novos nós se conectam de preferência a nós já altamente conectados confirmando a inovação como processo acumulativo e localizado. Além do mais, como discutido no capítulo 2, a dinâmica e retroalimentativa entre o cluster principal de tecnologia da GPT e seus diversos setores de aplicação fornecem oportunidades crescentes de inovação.

4.4. Considerações finais.

O presente trabalho propõe que a *World Wide Web* seja entendida como uma *General Purpose Technology*, uma tecnologia de aplicação generalizada e capaz de estimular a inovação nos mais variados setores da economia, ou complementariedades inovativas. A difusão desse conjunto de novas tecnologias interrelacionadas se dá pelos numerosos mecanismos de interação e feedbacks retroalimentativos entre os diversos setores de atividades e trajetórias tecnológicas.

O modelo de GPT, desenvolvido inicialmente por Bresnahan e Trajtenberg (1992 e 1995) determina que uma inovação radical, ao mostrar seu potencial de aplicabilidade generalizada e de lucratividade, atrai tanto novos empreendedores quanto o capital financeiro. Assim, a invenção da *World Wide Web*, na década de 1990, deu início a uma era marcada por uma rede mundial de informações disponível a qualquer indivíduo, à conectividade generalizada e a era das plataformas.

CONCLUSÃO

Shane Greenstein em seu livro “*How the Internet Became Commercial*” afirma:

As economias modernas frequentemente transformam tecnologias de ponta em tecnologias amplamente utilizadas – do misterioso ao banal. A Internet já foi exótica para todos, exceto para um pequeno grupo de conhecedores, mas há muito tempo a tecnologia se espalhou para a maioria das residências e empresas. No processo de se tornar onipresente, transformou a forma como trabalhamos e vivemos – mudando a forma como os consumidores se comportam e alterando a forma como as empresas fornecem produtos e serviços (p. 3, Greenstein)¹¹⁶.

É sob a luz desse entendimento de que a Internet, o mundo conectado, se estabeleceu como a norma organizadora da economia, da sociedade e das instituições, com transformações profundas sobre cada uma destas esferas, bem como de suas interações, que o presente trabalho se debruçou sobre o que entende ser a tecnologia que, de fato, permitiu a difusão e universalização da Internet e a criação de uma era de cada vez mais conectividade, a saber, a *World Wide Web*.

Como foi visto, a WWW surge da junção de uma necessidade contextual de um sistema de gerenciamento e compartilhamento de informações, especificamente de pesquisas científica no âmbito da CERN, com as inspirações pessoais de Berners-Lee de criar uma rede não hierárquica de informações que se igualasse a uma biblioteca universal de toda a informação disponível, com acesso democrático para todos que tivessem interesse. Ademais, é fruto da capacidade de articulação, por parte de seu inventor e de futuros colaboradores, de diversas tecnologias já disponíveis, mas até então não pensadas em conjunto para a solução de vários problemas que se impunham: a necessidade de tornar o acesso à Internet mais democrático e generalizado, de criar um sistema universal de armazenamento e acesso à informação, de superar gargalos do sistema de comunicação vigente, entre outros.

A WWW foi a responsável por tornar o acesso à Internet mais fácil e generalizado, impulsionando toda uma onda de investimentos e inovações nas tecnologias computacionais, de informação, telecomunicações, semicondutores etc. Também é considerada a forma pela qual a Internet virou um modelo de negócios, com o surgimento de toda uma nova gama de empresas e setores de atividades baseada na Internet, como o *e-commerce*, por exemplo. Além

116 Traduzido do original em inglês.

disso, indiretamente, impactou o avanço nos mais diversos setores, seja proporcionando melhores ferramentas de trabalho e pesquisa, seja na própria forma de organização do trabalho.

A WWW parece ter inaugurado uma nova trajetória tecnológica, onde a heurística e as formas de solucionar os problemas se encontram baseados no princípio da conectividade. Sendo assim, o presente trabalho se propôs a investigar a capacidade de transformação dessa tecnologia, não apenas em seus aspectos econômicos, através da utilização do conceito de *General Purpose Technologies*. Tal conceito, elaborado para enriquecer as discussões sobre revoluções tecnológicas, pode ser aplicado para entender tecnologias que causaram uma profunda mudança na economia, sociedade e nas instituições, mesmo que não alcancem o status de tecnologia deflagradora de *big-bang* no princípio de uma revolução tecnológica.

Para tanto, baseado na teoria de redes para análise de sistemas complexos e a partir da construção de duas redes, uma de patentes e outra de artigos, cujo primeiro nó é o artigo onde foi divulgada a tecnologia da WWW, procede-se a análise do comportamento das redes para identificação de suas principais características e estruturas que permitissem alguma identificação com as definições de tecnologias do tipo GPT. O que se observou foi a presença de redes com comportamentos típicos de redes tecnológicas e científicas, do tipo lei de potência, onde a presença de poucos *hubs* determinam em grande parte o comportamento da rede. Tais *hubs* sendo identificados com os principais setores tecnológicos de suporte e manutenção da WWW. Bem como um padrão de generalidade e pervasividade da tecnologia dada a propriedade *small world*. Através dessas e de outras referências e identificações, pode-se reforçar o argumento de ser a WWW uma GPT.

Ademais, o momento histórico atual é um reforço de como nossas vidas estão organizadas e baseadas em torno da Internet, assistimos, nos últimos anos, eleições cujos debates e resultados se deram praticamente na esfera das redes sociais, escândalos de espionagem internacional de governos de estados nacionais via coleta e análise de dados nas redes, vazamento de dossiês sobre dados de investimento em paraísos fiscais, debates sobre a “uberização do trabalho” com a crescente perda de direitos trabalhistas e precarização dos postos de trabalho, o surgimento de bolhas especulativas com moeda digitais e o debate sobre os impactos ambientais advindos da mineração desse tipo de ativo, NFTs e Metaverso. Por outro lado, diante de uma pandemia, foi possível manter empregos remotamente, escolas e tantas outras atividades. Lançando discussões ainda iniciais, por exemplo, sobre mudanças na forma de trabalho. Esses são alguns

dos milhares de exemplos que poderiam ser enumerados para demonstrar como a WWW alterou definitivamente a forma como produzimos e inovamos, a forma na qual o trabalho se organiza e em última instância a forma como vivemos.

REFERÊNCIAS BIBLIOGRÁFICAS:

ABBATE, Janet. **Inventing the Internet**. London: The MIT Press, 1999.

ALBUQUERQUE, E. Capitalismo pós-www: uma discussão introdutória sobre uma nova fase na economia global. **CADERNOS do DESENVOLVIMENTO**, Rio de Janeiro, vol. 14, n. 24, p. 131-154, jan.-jun. 2019.

ALBUQUERQUE, Eduardo da Motta. Revoluções tecnológicas e general purpose technologies: mudança técnica, dinâmica e transformações do capitalismo. In: RAPINI, Márcia Siqueira; RUFFONI, Janaina; SILVA, Leandro Alves; ALBUQUERQUE, Eduardo da Motta (org.). **Economia da ciência, tecnologia e inovação: fundamentos teóricos e a economia global**. 2.ed. Belo Horizonte: FACE – UFMG, 2021.

ARTHUR, W. Brian. **The Nature of Technology**. London: Penguin, 2009.

BEKAR, C.; CARLAW, K.; LIPSEY, R. General purpose technologies in theory, application and controversy: a review. **Journal of Evolutionary Economics**, n. 28, p. 1005–1033, 2018.

BERNERS-LEE, Tim; CAILLIAU, Robert; GROFF, Jean-François; POLLERMANN, Bernd. World-Wide Web: The Information Universe. **ELECTRONIC NETWORKING**, Vol. 2, No. 1, Spring 1992, pp. 52-58.

BERNERS-LEE, Tim; FISCHETTI, Mark. **Weaving the Web: the original design of the World Wide Web by its inventor**. New York: HarperCollins Publishers, 2000.

BOYACK, Kevin W.; KLAVANS, Richard. Creation and analysis of large-scale bibliometric networks. In: GLÄNZEL, Wolfgang; MOED, Henk F.; SCHMOCH, Ulrich; THELWALL, Mike (Eds.). **Springer Handbook of Science and Technology Indicators**. Springer Nature Switzerland, 2019.

BRESNAHAN, Timothy. General purpose technologies. In: HALL, Bronwyn H.; ROSENBERG, Nathan. **Handbook of the Economics of Innovation**. UK: North-Holland, 2010.

BRESNAHAN, T.; TRAJTENBERG, M. General purpose technologies: Engines of growth?, **NBER working paper**, n. 4148, 1992.

BRESNAHAN, T.; TRAJTENBERG, M. General purpose technologies: 'engines of growth'? **Journal of Econometrics**, vol. 65, n. 1, p. 83-108, 1995.

CABALLERO, Ricardo J.; JAFFE, Adam B. How High Are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth. In: JAFFE, Adam B.; TRAJTENBERG, Manuel. **Patents, citations, and innovations: a window on the knowledge economy**. Cambridge: Massachusetts Institute of Technology Press, 2002.

CERN. <https://home.cern/science/computing/birth-web>

DERTOUZOS, Michael L. **What Will Be**: How the new world of information will change our lives. London: Piatkus, 1997.

DOSI, G. Technological paradigms and technological trajectories. A suggested interpretation of the determinants and directions of technical change. **Research Policy**, n.11, p.147-162, 1982.

DOSI, Giovanni. **Technical Change and Industrial Transformation**. London: The Macmillan Press, 1984.

DOSI, G. Sources, procedures and microeconomic effects of innovation. **Journal of Economic Literature**, 26(3), 1120–1171, (1988a).

DOSI, G. The nature of the innovation process. In: DOSI, G.; FREEMAN, C.; NELSON, R. R.; SILVERBERG, G.; SOETE, L.:(eds). **Technical Change and Economic Theory**. Wheatsheaf: Brighton, 1988.

DOSI, Giovanni; ORSENIGO, Luigi. Coordination and transformation: an overview of structures, behaviours and change in evolutionary environments. In: DOSI, G.; FREEMAN, C.;

NELSON, R. R.; SILVERBERG, G.; SOETE, L.:(eds). **Technical Change and Economic Theory**. Wheatsheaf: Brighton, 1988.

DOSI, Giovanni; ORSENIGO, Luigi; LABINI, Mauro Sylos. Technology and the Economy. In: SMELSER, Neil J.; SWEDBERG, Richard (ed.). **The Handbook of Economic Sociology**. New Jersey: Princeton University Press, 2005.

FREEMAN, C.; LOUÇÃ, F. **As time goes by: from the industrial revolutions and to the information revolution**. Oxford: Oxford University, 2001

FREEMAN, C.; PEREZ, C. Structural crisis of adjustment: business cycles and investment behaviour. In: DOSI, G.; FREEMAN, C.; NELSON, R.; SILVERBERG, G.; SOETE, L. (eds). **Technical change and economic theory**. London: Pinter, 1988, p. 38-66.

GILLIES, James; CAILLIAU, Robert. **How the Web was born**. New York: Oxford University Press, 2000.

GLEICK, James. **The information: a history, a theory, a flood**. New York: Pantheon Books, 2011.

GREENSTEIN, S. **How the internet became commercial: innovation, privatization and the birth of a new network**. Princeton: Princeton University Press, 2015.

GRILICHES, Zvi, (ed). **R&D, Patents and Productivity**. University of Chicago Press, 1984.

GRILICHES, Zvi. Patent Statistics as Economic Indicators: a survey. **Journal of Economic Literature**, vol. 28, no. 4, p. 1661–1707, 1990.

HAFNER, Katie; LYON, Matthew. **Where Wizards Stay Up Late: The Origins of the Internet**. New York: Touchstone Edition, 1998.

HALL, Bronwyn H.; JAFFE, Adam B.; TRAJTENBERG, Manuel. The NBER Patent-Citations Data File: Lessons, Insights, and Methodological Tools. In: JAFFE, Adam B.;

TRAJTENBERG, Manuel. **Patents, citations, and innovations: a window on the knowledge economy**. Cambridge: Massachusetts Institute of Technology Press, 2002.

HELPMAN, Elhanan (ed). **General Purpose Technologies and Economic Growth**. Cambridge: Massachusetts Institute of Technology press, 1998.

HENDERSON, Rebecca; JAFFE, Adam B.; TRAJTENBERG, Manuel. Universities as a Source of Commercial Technology: A Detailed Analysis of University Patenting, 1965–1988. In: JAFFE, Adam B.; TRAJTENBERG, Manuel. **Patents, citations, and innovations: a window on the knowledge economy**. Cambridge: Massachusetts Institute of Technology Press, 2002.

KUHN, Thomas S. A estrutura das revoluções científicas. 13. ed. São Paulo: Perspectiva, 2017.

JAFFE, Adam B.; TRAJTENBERG, Manuel. **Patents, citations, and innovations: a window on the knowledge economy**. Cambridge: Massachusetts Institute of Technology Press, 2002.

JAFFE, Adam B.; TRAJTENBERG, Manuel, HERNERSON, Rebecca. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. In: JAFFE, Adam B.; TRAJTENBERG, Manuel. **Patents, citations, and innovations: a window on the knowledge economy**. Cambridge: Massachusetts Institute of Technology Press, 2002.

LIPSEY, Richard G.; BEKAR, Cliff; CARLAW, Kenneth. What requires explanation? In: HELPMAN, Elhanan (ed). **General Purpose Technologies and Economic Growth**. Cambridge: Massachusetts Institute of Technology press, 1998.

LIPSEY, R.G., CARLAW, K., BEKAR, C. **Economic Transformations: General Purpose Technologies and Long-Term Economic Growth**. Oxford University Press, Oxford; New York, 2005.

NAGAOKA, Sadao; MOTOHASHI, Kazuyuki; GOTO, Akira. Patent Statistics as an Innovation Indicator. In: HALL, Bronwyn H.; ROSENBERG, Nathan. **Handbook of the Economics of Innovation**. UK: North-Holland, 2010.

NELSON, Richard R. (ed.). **The Rate and Direction of Inventive Activity: Economic and Social Factors**. Princeton University Press., 1962.

NELSON, Richard R.; WINTER, Sidney G. In search of useful theory of innovation. **Research Policy**, 6, 36-76, 1977.

PEREZ, C. **Technological Revolutions and Financial Capital: the Dynamics of Bubbles and Golden Ages**. Cheltenham: Elgar, 2002.

PEREZ, C. Finance and Technical Change: A Neo-Schumpeterian Perspective. **CERF Working Papers**, n. 14, 2004.

PEREZ, C. Technological revolutions and techno-economic paradigms. **Cambridge Journal of Economics**, vol. 34, n. 1, p. 185-202, 2010.

PEREZ, C. Capitalism, technology and a green global golden age: the role of history in helping to shape the future. In: JACOBS, Michael; MAZZUCATO, Mariana (Eds.). **Rethinking capitalism: economics and policy for sustainable and inclusive growth**. Chichester, West Sussex, United Kingdom: Wiley-Blackwell, in association with The Political Quarterly, 2016.

ROSENBERG, Nathan. **Perspectives on technology**. New York: Cambridge University Press 1976.

ROSENBERG, N. Chemical engineering as a General Purpose Technology. In: HELPMAN, E. **General Purpose Technologies and economic growth**. Cambridge, Mass./London: The MIT Press, p. 167-192, 1998.

ROSENBERG, Nathan. **Por dentro da caixa-preta: tecnologia e economia**. Campinas, SP: Editora Unicamp, 2006.

SCHMOOKLER, Jacob. **Invention and Economic Growth**. Harvard University Press, 1966.

SCHUMPETER, J. (1911) **A teoria do desenvolvimento econômico**. São Paulo: Nova Cultural, 1985.

SCHUMPETER, J. (1939) **Business cycles: a theoretical, historical and statistical analysis of the capitalist process**. Philadelphia: Porcupine, 1989.

SCHUMPETER, J. (1942) **Capitalismo, socialismo e democracia**. Rio de Janeiro: Zahar Editores, 1984.

SONI, Jimmy; GOODMAN, Rob. **A mind at play: How Claude Shannon Invented the Information Age**. New York: Simon & Schuster, 2017.

SRNICEK, Nick. **Platform capitalism**. Cambridge, UK; Malden, MA: Polity Press, 2016.

STEPHAN, Paula E. The Economics of Science. In: HALL, Bronwyn H.; ROSENBERG, Nathan. **Handbook of the Economics of Innovation**. UK: North-Holland, 2010.

TRAJTENBERG, Manuel. **Patents, citations, and innovations: a window on the knowledge economy**. Cambridge: Massachusetts Institute of Technology Press, 2002.

TRAJTENBERG, Manuel. A Penny for Your Quotes: Patent Citations and the Value of Innovations. In: JAFFE, Adam B.; TRAJTENBERG, Manuel. **Patents, citations, and innovations: a window on the knowledge economy**. Cambridge: Massachusetts Institute of Technology Press, 2002.

TRAJTENBERG, Manuel; HENDERSON, Rebecca; JAFFE, Adam B. University versus Corporate Patents: A Window on the Basicness of Invention. In: JAFFE, Adam B.; TRAJTENBERG, Manuel. **Patents, citations, and innovations: a window on the knowledge economy**. Cambridge: Massachusetts Institute of Technology Press, 2002.

VAN RAAN, Anthony. Measuring science: basic principles and application of advanced bibliometrics. In: GLÄNZEL, Wolfgang; MOED, Henk F.; SCHMOCH, Ulrich; THELWALL, Mike (Eds.). **Springer Handbook of Science and Technology Indicators**. Springer Nature Switzerland, 2019.

VON TUNZELMANN, Nick; MALERBA, Franco; NIGHTINGALE, Paul; METCALFE, Stan. Technological paradigms: past, present and future. **Industrial and Corporate Change**, Volume 17, Number 3, pp. 467–484, 2008.

YOFFIE, David B. **Competing in the age of digital convergence**. Boston: Harvard Business School Press, 1997.

ZUBOFF, Shoshana. **The age of surveillance capitalism: the fight for a human future at the new frontier of power**. New York: PublicAffairs, 2018.