

**SUZANE MANGINI ALMEIDA**

**USO DE BIG DATA EM MÍDIAS SOCIAIS: PANORAMA  
ATUAL NA CIÊNCIA**

**PUC-CAMPINAS**

**2018**

**SUZANE MANGINI ALMEIDA**

**USO DE BIG DATA EM MÍDIAS SOCIAIS: PANORAMA  
ATUAL NA CIÊNCIA**

Dissertação apresentada como exigência para conclusão do Mestrado Interdisciplinar em Linguagens, Mídia e Arte, do Centro de Linguagem e Comunicação, da Pontifícia Universidade Católica de Campinas.

Orientador: Prof. Dr. César Antônio Pereira

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DE CAMPINAS**

**CAMPINAS - 2018**

Ficha catalográfica elaborada por Marluce Barbosa – CRB 8/7313  
Sistemas de Bibliotecas e Informação – SBI – PUC-Campinas

t001.5  
A447u Almeida, Suzane Mangini.  
Uso do Big Data em mídias sociais: panorama atual da ciência /  
Suzane Mangini Almeida. - Campinas: PUC-Campinas, 2018.  
76f.

Orientador: César Antonio Pereira.  
Dissertação (mestrado) - Pontifícia Universidade Católica de Campi-  
nas, Centro de Linguagem e Comunicação, Pós-Graduação em Lingua-  
gens, Mídia e Arte.  
Inclui bibliografia.

1. Análise de conteúdo (Comunicação). 2. Mídia social. 3. Big data. 4.  
Bibliometria. I. Pereira, César Antonio. II. Pontifícia Universidade Católica  
de Campinas. Centro de Linguagem e Comunicação. Pós-Graduação em  
Linguagens, Mídia e Arte. III. Título.

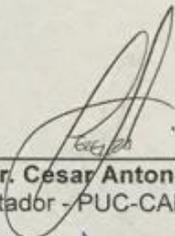
22. ed. CDD – t001.5

SUZANE MANGINI ALMEIDA

"USO DE BIG DATA EM MÍDIAS SOCIAIS: PANORAMA ATUAL NA CIÊNCIA"

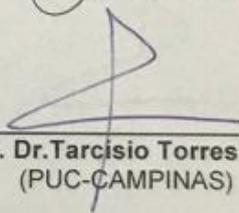
Este exemplar corresponde à redação final da Dissertação de Mestrado em Linguagens, Mídia e Arte da PUC-Campinas, e aprovada pela Banca Examinadora.

APROVADA: 19 de dezembro de 2018.



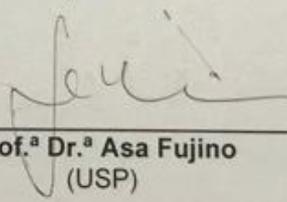
---

**Prof. Dr. Cesar Antonio Pereira**  
(Orientador - PUC-CAMPINAS)



---

**Prof. Dr. Tarcisio Torres Silva**  
(PUC-CAMPINAS)



---

**Prof.ª Dr.ª Asa Fujino**  
(USP)

“O começo de todas as ciências é o espanto  
de as coisas serem o que são”.

Aristóteles

## **AGRADECIMENTOS**

Primeiramente agradeço a Deus pelas “coincidências” que me fizeram chegar até aqui. Pela oportunidade e por me capacitar a perseguir nesta jornada de conhecimento e aprendizagem.

Aos meus pais, pelo incentivo e apoio sempre presente em toda minha trajetória como estudante e principalmente em todos os desafios e conquistas da minha vida. Vocês são meu porto seguro e minha referência.

À toda a minha família e amigos, pelo estímulo e compreensão aos encontros mais curtos e escassos.

Ao meu noivo, pelo incentivo, paciência e carinho durante toda essa jornada.

Aos meus chefes e colegas, pela tolerância e flexibilidade de horários, permitindo que fosse possível desempenhar ambas as atividades.

Aos meus colegas de curso, que enriqueçam meu olhar com diferentes perspectivas de projetos, compreensões dos autores e infinitos apoios. Estar ao lado de vocês fez desse processo muito mais rico e prazeroso.

À secretaria do programa, pela atenção em cada detalhe e apoio constante.

Aos professores da PUC Campinas que me acolheram, me apoiaram e me apresentaram um mundo de possibilidades. Em especial meu querido orientador, César, pelo cuidado, flexibilidade e pela doação de tempo e conhecimentos. Você foi peça fundamental para que esse trabalho fosse concluído.

À banca, pela contribuição de forma tão rica na construção desse estudo.

Todos vocês aceitaram e compartilharam do meu sonho, contribuindo para a sua realização. Sou muito grata por esta oportunidade e por essa grande rede de apoio e carinho que me cerca.

## RESUMO

Presente às transformações impulsionadas pela internet, verificou-se a disseminação do uso de mídias sociais e o emprego do Big Data como ferramenta capaz de estabelecer relação entre uma grande quantidade de dados inaugurando um campus de possibilidades para a pesquisa científica nos mais diversos campos do conhecimento.

A pesquisa apresentou um panorama da produção científica que relaciona o uso de Big Data e Mídias Sociais segundo artigos publicados em periódicos indexados na base de dados *Web of Science*. Com caráter exploratório e de abordagem qualitativa-quantitativa, o estudo aplicou procedimentos de análises bibliométrica e cientométrica, além de análise de conteúdo. O corpus constituiu-se de 541 artigos que permitiu identificar o comportamento temático e empírico a luz do posicionamento geográfico e temporal, bem com analisar as contribuições e limitações nos diferentes campos do conhecimento.

Os resultados evidenciam ampla diversificação de campos do conhecimento com interesses de pesquisa sobre a temática distribuídos em curto espaço de tempo, em variados contextos temáticos e em crescimento quantitativo. Com relação ao posicionamento geográfico identifica-se grande interesse de pesquisa em muitos países com grande concentração nos Estados Unidos da América. Fundamentalmente as pesquisas sobre Big Data e mídia social encontram-se em processo de consolidação na ciência.

**Palavra-chave:** Big Data, mídia social, cientometria, bibliometria

## **ABSTRACT**

Face to the transformations driven by the internet, it was verified the dissemination of social media usage and the application of Big Data as a tool capable of establishing a relationship between a great amount of data inaugurating a campus of possibilities for scientific research in the most diverse fields of knowledge .

The research presented an overview of the scientific production that relates the use of Big Data and Social Media according to articles published in journals indexed in the Web of Science database. With an exploratory and qualitative-quantitative approach, the study applied bibliometric and scientometric analysis procedures, as well as content analysis. The corpus consisted of 541 articles that allowed the identification of thematic and empirical behavior in the light of the geographical and temporal positioning, as well as analyzing the contributions and limitations in the different fields of knowledge.

The results show a wide diversification of fields of knowledge with research interests on the subject distributed in a short period of time, in various thematic contexts and in quantitative growth. Regarding geographic positioning, there is great research interest in many countries with great concentration in the United States of America. Fundamentally research on Big Data and social media is in the process of consolidation in science.

**Key words:** Big data, social media, bibliometric, scientometric

## LISTA DE GRÁFICOS

Gráfico 1. Número de artigos por ano e taxa de crescimento.....	46
Gráfico 2. Número de artigos por ano sobrepostos Big Data, Mídias Sociais e sua combinação .....	47
Gráfico 3. Número de publicações por área de conhecimento (17 áreas do conhecimento mais expressivas).....	49
Gráfico 4. Distribuição geográfica dos artigos excluindo Estados Unidos .....	50
Gráfico 5. Diagrama de correlação de áreas do conhecimento.....	51
Gráfico 6. Diagrama de correlação de ano de publicação e área do conhecimento.....	53
Gráfico 7. Wordcloud principais palavras chaves .....	54
Gráfico 8. Diagrama de correlação de ano de publicação e palavra-chave .....	59
Gráfico 9. Ano de publicação das principais referências .....	63
Gráfico 10. Número de citações por ano .....	64

## LISTA DE TABELAS

Tabela 1. Principais mídias sociais do mundo.....	30
Tabela 2. Evolução da utilização dos dados e informação.....	33
Tabela 3. Periódicos mais representativos para a temática .....	48
Tabela 4. Palavras-chave por área do conhecimento .....	56
Tabela 5. Palavra-chave por geografia.....	58
Tabela 6. Autores mais representativos para a temática.....	62
Tabela 7. Instituições mais representativas para a temática .....	65

## SUMÁRIO

1. INTRODUÇÃO .....	11
2. REFERENCIAL TEÓRICO.....	17
2.1 SOCIEDADE HIPERCONECTADA .....	17
2.2 MUDIATIZAÇÃO E MÍDIA SOCIAL.....	25
2.3 BIG DATA, MUITO ALÉM DE BANCO DE DADOS .....	32
3. METODOLOGIA .....	40
3.1. ESTRATÉGIA DE COLETA DOS DADOS .....	42
3.2. PROCEDIMENTOS PARA ANÁLISE .....	44
4. RESULTADOS E DISCUSSÃO .....	45
5. CONCLUSÃO E CONSIDERAÇÕES FINAIS .....	67
6. REFERÊNCIAS .....	72

## 1. INTRODUÇÃO

Fruto do acelerado desenvolvimento tecnológico, impulsionado pela internet, o período atual é marcado por rápidas e significativas mudanças que reverberam em diferentes áreas. (CASTELLS, 1999 e ZALIS, 2016). O uso difundido de aparelhos, como notebooks e smartphones, conectados em rede aumentaram a quantidade e a velocidade da produção, coleta e uso de informações, que conferiu ao atual momento a denominação de *Era da Informação* (RÜDIGER e PIMENTA, 2013)

Esta evolução tecnológica reflete-se na comunicação, que com o desenvolvimento das primeiras Tecnologias de Informação e Comunicação, TICs, como rádio e televisão, começou a ser comumente empregado o termo mídia para designar todos os meios de comunicação. O uso da internet propiciou a expansão de novos canais de relacionamentos e interconexões sociais, redes sociais (BRATTON e EVANS, 2012; EVANS e MCKEE, 2010; WIKSTRÖM e WIGMO, 2010). Com o surgimento e disseminação das redes sociais como uma forma de comunicação, essas passaram a ser comumente conhecidas como mídias sociais. (SANTAELLA, 2003).

Diferentemente das mídias de massa tradicionais como TV e rádio, o uso da internet permitiu que todo usuário, além de receptor, fosse também produtor, manifestando suas percepções e produzindo conteúdos, ou seja, ampliando o fluxo de comunicação, permitindo que ele ocorra em ambos os sentidos. Além disso, dada a possibilidade de contato em rede, as interações que ocorrem utilizando-se da internet se dão em tempo real e com conexões mundiais, provocando o crescente aumento no número da produção e compartilhamento de informações no formato de textos, imagens, vídeos e outros, que quando disseminados, estimulam mais interações. (RÜDIGER e PIMENTA, 2013).

Devido a sua natureza digital, o que é feito e comunicado utilizando a internet não passa despercebido e tão pouco é esquecido. Cada email recebido, música ou filme consumido ou URL consultada é gravado digitalmente nos servidores que hospedam a web. Desta forma, todos os movimentos, cliques e postagens realizadas no ciberespaço podem ser registrados e armazenados em forma de bancos de dados. (RÜDIGER e PIMENTA, 2013). Estes registros possibilitam a personalização de navegação e outras facilidades como detectar um spam, oferecer recomendações e anúncios segmentados. (GOLDER e MACY, 2012). De forma similar, os mesmos traços de atividades que tornam

possíveis as mídias sociais de oferecerem serviços funcionais para os seus usuários, também podem ser usados por cientistas sociais para diretamente observar o comportamento humano em detalhes e como se desdobra ao longo do tempo em uma escala global. (GOLDER e MACY, 2012).

Compreendendo que é a sociedade que ajusta a tecnologia às suas necessidades, portanto, que a projeta e não o oposto, ao estudar as tecnologias e principalmente as transformações que a permeiam, é possível melhor compreender os comportamentos e valores da sociedade que a detém (CASTELLS, 1999). Quando observado o uso e interações dos usuários em mídias sociais, Golder e Macy (2012) destacam que embora existam diferenças da vida social offline, o online reflete as mesmas buscas, dúvidas, pensamentos e relações do usuário. Desta forma, é possível identificar e estudar a sociedade ao melhor compreender sua relação com a tecnologia e o uso de mídias sociais, o que ganha relevância a medida que permite o acesso a uma grande quantidade de registros digitais quando as relações são permeadas por tecnologia. A utilização de grandes quantidades de traços digitais do comportamento humano, representa um novo paradigma de coleta de dados para a pesquisa em Ciências Sociais e para a ciência como um todo. (GOLDER e MACY, 2012)

Uma vez que o aprimoramento da ciência se dá a partir da coleta e observação dos acontecimentos, nos seus primórdios as pesquisas se baseavam em ensaios feitos em laboratórios e dados geralmente disponíveis em cadernos científicos pessoais (Gordon Bell, 2011). Dada a evolução tecnológica houve um aprimoramento nos instrumentos de observação como, por exemplo, telescópios, satélites e sensores especializados que ampliaram as possibilidades de pesquisa e permitiram a coleta de dados mais ampla e precisa. Devido à variedade de tópicos, em grande escala, e a possibilidade de análise por variados períodos de tempo, as mídias sociais proporcionam grande quantidade de material de pesquisa que despertam o interesse de pesquisadores de diferentes áreas para entender melhor as relações sociais e o comportamento (Brickman Bhutta 2012; Golder and Macy 2011; Heavilin et al. 2011; Lowe et al. 2012; Moreno et al. 2012; Valkenburg, Peter, and Schouten 2006).

Compreendendo a produção de informação como o processo de reunião, seleção, codificação, armazenamento, ordenação e estoque da informação e considerando que a informação precisa ser compreendida para então ser transformada em conhecimento capaz de ser utilizado (BARRETO, 1998), entende-se que as TICs, advindas da internet,

em especial as mídias sociais, estimulam a produção de informação, que oferece crescente potencial de ser transmitida e compreendida, ou seja, tornar-se conhecimento, passível de ser utilizado em variados contextos, sobretudo na ciência, no aprimoramento das tecnologias existentes e produzindo inovações.

Desta forma, há na utilização dos dados gerados a partir do uso de mídias sociais rico potencial informacional capaz de promover o diferencial competitivo necessário para a perpetuação das iniciativas diante do presente momento de elevado ritmo de inovação e competição para variados atores sociais. As empresas e instituições precisam acompanhar o dinamismo vigente para empregar seus recursos de forma mais eficiente e ágil, em prol dos resultados desejados (LICKS, 2016). Desponta, portanto, não só a necessidade de ordenação dos dados e das informações continuamente produzidas, mas principalmente a geração de insumos informacionais que possam contribuir para a tomada de decisões.

Sendo assim, para possibilitar os estudos desejados faz-se necessária a disponibilização e ordenação dos dados coletados, sobretudo devido à quantidade de compartilhamentos em nível global e a velocidade de geração de informação, que aumenta progressivamente, as técnicas tradicionais utilizadas para sua análise e armazenamento revelam-se obsoletas, urgindo a necessidade de desenvolver padrões mais modernos de organização, a fim de disponibilizar a informação (CALDAS e SILVA, 2016).

Neste cenário emergiu o papel do *Big Data*, expressão que se refere à manipulação de grande quantidade de dados, para se compreender, prever e analisar algo que não seria possível através de uma pequena amostragem. O seu diferencial é relacionar os dados de forma a apresentar uma interação entre eles e não uma análise de sua causa. Sendo assim, o tratamento desses dados auxilia a pontar a ocorrência estatística de um fato fornecendo insumos para projeções, descobertas e plano ações de forma mais estratégica. (CUKIER & MAYER-SCHOENBERGER, 2013 e OLIVEIRA, 2015).

Sendo o Big Data uma ferramenta capaz de encontrar relacionamento entre diferentes fontes de informações e analisar grande quantidade de dados, quando inserido em plataformas sociais auxilia no entendimento de diversos aspectos de interações humanas, o que confere a seu uso um grande valor. Somando-se a sua utilidade, quando empregado às mídias sociais é possível desfrutar de condições

bastante favoráveis à sua melhor utilização pelos seguintes fatores: nestas plataformas os dados estão abertamente disponíveis e sem custos, sua coleta é relativamente simples, sua cobertura espacial é bastante grande (local e global), os dados são continuamente (e rapidamente) gerados e o seu conteúdo é rico. (TENKANEN, MININ, HEIKINHEIMO, HAUSMANN, HERBST et. Al., 2017)

Uma vez que as informações e dados gerados e compartilhados em rede podem ser das mais diferentes naturezas (imagem, vídeo, áudio e texto), e a coleta está amplamente diversificada, sua utilização e importância também está disseminada nos mais diversos setores e, a cada dia afeta mais os envolvidos. Com aplicações em diferentes setores, como da saúde, conferindo maior precisão nos diagnósticos, auxiliando na gestão pública, na educação, nos transportes e mobilidade urbanas, nas redes varejistas e inclusive no meio acadêmico, que se utiliza dos recursos para análise científica ao mesmo tempo que se utiliza para melhor disponibilização da produção científica, dentre outros. Enfim, se expande para todos os campos do conhecimento humano. (CALDAS e SILVA, 2016).

Dentre as várias vantagens apontadas pelo uso de Big Data destaca-se a oportunidade de redução de custos, agilidade no processamento de tarefas e até mesmo a capacidade tornar possível à execução de tarefas e análises sem o qual não seria factível devido à quantidade de tempo e recursos que demandariam ser utilizados. Desta forma, o uso do Big Data oferece maior respaldo para a tomada de decisão, seja em qual for o setor que está endereçando. (LICKS, 2016). E torna-se, portanto, base fundamental para o desenvolvimento da ciência e, por conseguinte de inovações, que por si carregam o potencial de desenvolvimento econômico e melhores condições de vida, além de benefícios ao meio ambiente como um todo.

Ao mesmo tempo que a análise de dados provenientes das mídias sociais despertam o interesse e podem ser utilizados em ações de melhoria para a sociedade, vê-se surgir estudos que alertam sobre o seu uso excessivo e possíveis usos inapropriados, bem como reflexões sobre a aplicação de análise de dados para determinar algum tipo de controle dos usuários e suas opiniões. Esse último evidenciado por conta das polêmicas vinculadas às estratégias eleitorais de diversos países. Conforme observado por Hemsley (2018) a mídia pode ser usada tanto para o bem social, ao possibilitar derrubar regimes opressivos ou trazendo desigualdade de renda à discussão, ao mesmo tempo em que sempre haverá aqueles que aproveitam a

tecnologia de maneira que não foram intencionadas pelos designers como com o uso de notícias falsas e direcionamento de algoritmos entre outros.

Boyd e Crawford (2012) aponta o questionamento se de fato o desejado acesso a quantidades massivas de informações produzidas por e sobre pessoas, coisas e suas interações ajudarão a criar melhores ferramentas, serviços e bens públicos ou inaugurar uma nova onda de incursões de privacidade e marketing invasivo. Se a análise de dados ajudará a entender comunidades on-line e movimentos políticos ou será usado para rastrear os manifestantes e suprimir a fala. Desta forma, os autores reforçam a necessidade de interrogar criticamente suas suposições e vieses, dada a ascensão do Big Data como um fenômeno sócio-técnico. Visto que consideram as questões do Big Data um fenômeno cultural, tecnológico e acadêmico que provoca extensa retórica utópica e distópica.

Apesar das possibilidades para as pesquisas sociais provenientes do uso dos dados das mídias sociais e da necessidade de uma melhor análise e reflexão no que tange o uso destes recursos, Golder e Macy (2012, p. 7) apontam que os cientistas sociais costumam ver estes dados como inacessíveis para pesquisa em ciências sociais e unicamente relevantes para ciência da computação e outras áreas afins. Os autores lamentam, "a maioria da ciência social e comportamental usando dados on-line está vindo de da computação e cientistas da informação que nem sempre têm o treinamento necessário para perguntar as questões certas, ou para reconhecer pressupostos infundados e ramificações socialmente injustas." Desta forma, os autores evidenciam a necessidade de se confirmar se os dados e recursos científicos provenientes das mídias sociais estão de fato sendo aproveitados por todas as áreas do conhecimento de acordo com sua máxima possibilidade de aproveitamento.

Portanto, percebe-se necessário o conhecimento sobre as contribuições epistemológicas e empíricas proporcionadas pelo uso do *Big Data* em mídias sociais que possibilitem compreender o estado atual de aplicação e de uso em diferentes campos do conhecimento, oportunidades e desafios. A carência na compreensão dessas estruturas dificulta o entendimento do estágio de amadurecimento alcançado, inibe o aperfeiçoamento teórico e prático do conhecimento produzido em seus variados espaços e inviabiliza projeções e avanços à luz das demandas científicas. Os marcos temporais, geográficos, metodológicos e teóricos que regulam a produção do conhecimento sobre

a temática tornam-se incompreensíveis e desconexos à realidade. Consequentemente, há prejuízos no desenvolvimento da avaliação da atividade científica.

Logo, parte-se da hipótese da crescente utilização e aplicação de Big Data em mídias sociais confirmando-se a possibilidade de emprego sob diversas perspectivas temáticas e permitindo o desenvolvimento e aprimoramento de técnicas de utilização na obtenção de conteúdo informacional para a ciência. A temática está em processo de consolidação na ciência passível de ser verificado a partir de crescentes estudos desenvolvidos em diversificadas áreas do conhecimento, por múltiplos pesquisadores e instituições, além de estar geograficamente distribuída com diferentes abordagens temáticas. Quanto a isso, pressupõem-se grande impacto de estudos desenvolvidos a partir da lógica de aplicação da ferramenta de Big Data em mídia social, além de marcos teóricos pouco desenvolvidos, ancorados em poucas áreas do conhecimento, sobretudo nas ciências exatas. Portanto, os estudos sobre o uso de Big Data em mídia social na ciência configuram-se em consolidação.

Assim sendo, esta pesquisa mostra-se relevante a medida que elucida o estudo sobre o uso de Big Data em mídia social refletido nos diferentes campos do conhecimento científico, apontando tendências, oportunidades e lacunas. Devido à efervescência e contemporaneidade do tema, acrescido seu potencial informacional acerca da sociedade e seus indivíduos, a compreensão de sua dinâmica atual permite visão ampla e interdisciplinar capaz de auxiliar no desenvolvimento de estratégias que visam aperfeiçoar os processos e estabelecer ações assertivas para obtenção de resultados de maior relevância e com maior agilidade aos processos científicos.

Se valendo destas considerações, o presente estudo tem por objetivo compreender como a ciência tem se valido acerca do uso de Big Data em mídias sociais, identificando objetos teóricos, empíricos e potenciais contribuições nos diferentes campos do conhecimento e em especial no que se refere ao enfoque da ciência social aplicada. Complementarmente, busca-se identificar o comportamento temático nos diferentes campos do conhecimento, tendo como referência os assuntos abordados e seu comportamento empírico à luz do posicionamento geográfico e temporal.

Por sua vez, a análise das contribuições e limitações da temática nos diferentes campos do conhecimento, vem a permitir a compreensão das carências e oportunidades de aprofundamentos teóricos e empíricos na ciência.

## **2. REFERENCIAL TEÓRICO**

### **2.1 SOCIEDADE HIPERCONECTADA**

Segundo Oliveira (2003), a produção científica está diretamente relacionada às demandas do momento histórico, assim como às possibilidades oferecidas ao seu desenvolvimento. Desta forma o conhecimento científico não pode ser dissociado da vida humana e da relação social, resultando da ação dialógica entre as complementariedades e antagonismos da razão, da experiência, da imaginação e da verificação. Fazendo-se necessária a contextualização do momento social para melhor compreensão do conhecimento científico produzido.

Nesse sentido, as últimas décadas foram marcadas por transformações que, em grande parte, foram atribuídas aos avanços tecnológicos associados com a propagação da internet. Mais do que uma rede digital de comunicação capaz de conectar todo o mundo sem limitação de localização e alcance, a internet tem sido considerada como o agente de mudanças de diversas áreas do cotidiano, como o trabalho, as relações sociais, afetivas e de identidade.

A exemplo dos novos produtos e serviços que se beneficiaram do avanço da internet, o lançamento do iPhone em 2007 se destacou como um expoente de transformações. O aparelho ganhou importância não por ser o primeiro smartphone, mas por obter larga aceitação e por inaugurar um novo formato, abrindo caminho para outras inovações. O dispositivo, através de seus aplicativos, incorporou atividades que antes eram realizadas através de outros meios e tornou-se item indispensável, sendo incorporado como uma extensão do próprio corpo. Conforme Alves (2017, p.24) “A máquina, mais do que nunca, era uma extensão do nosso corpo; estaria talvez para sempre acoplada a nós”.

Levando-se em consideração que a sociedade ajusta a tecnologia às suas necessidades, portanto, que a projeta e não o oposto, ao estudar às tecnologias e principalmente as transformações que a permeiam, é possível melhor compreender os comportamentos e valores da sociedade que a detém (CASTELLS, 1999).

“Nós sabemos que a tecnologia não determina a sociedade: é a sociedade. A sociedade é que dá forma à tecnologia de acordo com suas

necessidades, valores e interesses das pessoas que utilizam as tecnologias. Além disso, as tecnologias de comunicação e informação são particularmente sensíveis aos efeitos dos usos sociais da própria tecnologia". (CASTELLS, 1999, p.25)

No entanto, por mais que a sociedade não determine a tecnologia, segundo Castells (1999), ela pode retardar ou acelerar seu processo de desenvolvimento. Desta forma, a tecnologia pode não determinar a evolução histórica e transformação social da sociedade, mas ela incorpora a capacidade de transformação das sociedades, bem como os usos que as sociedades decidem dar ao seu potencial tecnológico. Para o autor não há distinção nítida de tecnologia e sociedade, ao afirmar que são necessárias as ferramentas tecnológicas para que seja possível entender ou representar uma sociedade.

Assim, sendo as máquinas entendidas como atores e não como causa e efeito, podemos compreender a tecnologia não como determinante das mudanças, mas como agente, uma vez que participa dos processos sociais e culturais. Consequentemente, as formações de redes, estimuladas pelas condições tecnológicas desenvolvidas, são também fruto e catalisadores de uma nova forma de organização social, uma vez que estão inseridas em um conjunto de transformações que facilitam a comunicação e interação entre indivíduos, o que acelera a difusão de ideias que culminarão em inovações. Neste contexto, para melhor compreender os avanços tecnológicos experienciados e sua interrelação com os seres humanos, é preciso entender a dinâmica e anseios desta sociedade, e suas relações com os avanços tecnológicos.

A rede digital que se forma mediante o uso de internet, de acordo com Levy (1998), pode ser chamada de *ciberespaço* e é o meio de comunicação que emerge da interconexão dos computadores em nível global. Paralelamente, o conjunto de atitudes, técnicas e modos de agir neste espaço é resumidamente compreendido como *cibercultura*. Desta forma, utiliza-se estas nomenclaturas para melhor designar o meio e o conjunto de atitudes compartilhados em meio digital.

Diferentemente do que ocorre nos veículos de comunicação em massa, no qual a comunicação flui majoritariamente em um único sentido, fluxo que foi considerado por teorias de comunicação, no ciberespaço todo usuário além de receptor pode ser também produtor, revelando uma mudança na relação com a mídia e que permeia outras esferas do dia a dia. Para além do âmbito da comunicação, Toffler (1980) introduziu o termo

*prosumer* imaginando um tipo de consumidor do futuro, envolvido no desenho e na manufatura de produtos para torná-los mais personalizados e individualizados.

Corroborando, Kozinets (2007) definiu os *prosumers* como consumidores que se identificam como membros de determinado grupo e que coletivamente usam uma cultura de consumo - e esse uso inclui a construção individual e coletiva de práticas, identidades e significados (sobrepostos e mesmo conflitantes), textos, imagens e objetos alternados. Ou seja, o que se observa no momento atual é busca por consumos mais personalizados, com maior identificação e participação dos indivíduos, que, dentro do ciberespaço ganha força devido a sua natureza em rede e altamente ajustável.

Dado a sua natureza digital, todos os movimentos, cliques e postagens realizadas no ciberespaço podem ser registrados e armazenados na forma de bancos de dados. Em rede, as interações ocorrem em tempo real e com conexões mundiais. Desta forma, observa-se um número crescente na produção e compartilhamento de textos, imagens, vídeos e outros formatos, que quando disseminados, estimulam mais interações, que, por conseguinte, serão registradas na forma de banco de dados. (RÜDIGER e PIMENTA, 2013).

Este volume de iterações, que é acelerado pelo uso da internet, reflete-se no cotidiano da sociedade, proporcionando maior quantidade e acesso às informações em escala global, fazendo com que seja mais rápido estar atualizado, ao mesmo tempo que há uma sobrecarga e dificuldade de filtros, visto que há um volume muito grande de informações a serem absorvidas.

Transformações como estas, relacionadas aos avanços tecnológicos, são apontadas por estudiosos como características do momento atual em que impera a ordem das “mudanças sociais e culturais verdadeiramente importantes” iniciadas no final do século XX (LYON, 1998, p. 9). Este momento é conhecido como pós-modernidade, e em diferentes abordagens ganha denominações como: modernidade líquida, capitalismo tardio, hipermodernidade, modernidade reflexiva, entre outros, a depender da análise de cada autor por diferentes perspectivas acerca da economia, política ou sociedade.

Embora haja divergências de nomenclaturas e perspectivas, elas se aproximam no que se refere a importância das transformações observadas. Lyotard (1989), predizia o pós-modernismo, já na década de 1970, ao abordar o enfraquecimento das verdades e das leis, que centravam as questões do saber nas ciências desde o iluminismo.

Posteriormente, Jameson (2011) relacionou a pós-modernidade à globalização como agentes simultâneos das transformações observadas, considerando a globalização a infraestrutura da qual a pós-modernidade era a superestrutura, que, a partir dessa conexão, desencadearam as transformações marcadas pelas tecnologias da informação, pelas redes globais, pelo capital financeiro, pela descolonização e pela introdução do computador pessoal ao cotidiano (JAMESON, 2011).

Por sua vez, Baumaun (2001) utiliza-se do conceito de líquido para retratar o contexto atual repleto de instabilidades e associa à globalização em virtude desta acarretar processos de união e divisão, seja no campo político, econômico ou social. A característica disforme dos líquidos – eles assumem a forma do recipiente onde estão – se aplica aos indivíduos contemporâneos, que “não se atêm muito a qualquer forma e estão constantemente prontos (e propensos) a mudar” (BAUMAUN, 2001, p.8)

O abandono das formas sólidas da modernidade, conforme argumentado por Baumaun, sugere uma liberdade que foi almejada, mas que acompanhada da perda da segurança, tornou-se motivo de mal-estar. Baumaun (1998, p.10) afirma que “os homens e as mulheres pós-modernos trocaram um quinhão de suas possibilidades de segurança por um quinhão de felicidade”, embora não acredite que esta conquista tenha rendido satisfação plena.

[...] liberdade sem segurança não assegura mais firmemente uma provisão de felicidade do que segurança sem liberdade. Uma disposição diferente das questões humanas não é necessariamente um passo a diante no caminho da maior felicidade: só parece tal no momento em que se está fazendo. A reavaliação de todos os valores é um momento feliz, estimulante, mas os valores reavaliados não garantem necessariamente um estado de satisfação. (BAUMAUN, 1998, p.10)

Sendo assim, verifica-se a constância da incerteza como o único parâmetro de certeza (BAUMAN, 2011). O pode ser observado pela valorização do imediatismo, que, ao ser associado à possibilidade de acesso instantâneo, configura uma pressa no descarte e substituição, práticas que são facilitadas pela dinâmica do ciberespaço,

a velocidade atordoante da circulação, da reciclagem, do envelhecimento, do entulho e da substituição que traz lucro hoje – não a durabilidade e confiabilidade do produto. Numa notável reversão da tradição

milênar, são os grandes e poderosos que evitam o durável e desejam o transitório, enquanto os da base da pirâmide – contra todas as chances – lutam desesperadamente para fazer suas frágeis, mesquinhas e transitórias posses durarem mais tempo. Os dois se encontram hoje em dia principalmente nos lados opostos dos balcões das mega liquidações ou de vendas de carros usados. (BAUMAN, 2011, p.21)

Para além do conceito de instabilidade, Lipovetsky (2004) exalta a potência e o exagero presente no momento atual. Ele atualiza o conceito de pós-modernidade e chega à “hipermodernidade” – fase iniciada na década de 1990 e que se mantém pertinente até os dias de hoje em sua incompletude. A hiper circulação do capital financeiro, das informações por meio dos inúmeros avanços tecnológicos, principalmente, pela internet são instaurados na hipermodernidade.

O prefixo “hiper” remete a uma época de excessos, em que tudo é elevado a uma potência superlativa e essa é a ideia sugerida por Lipovestky (2004) quando questiona “o que mais não é *hiper*?”, hoje em dia, e apresenta uma lista de palavras – as quais estão frequentemente presentes em nossas vidas: “hipercapitalismo, hiperclasse, hiperpotência, hiperterrorismo, hiperindividualismo, hipermercado, hipertexto.” (LIPOVESTKY, 2004, p.52). A relação com o tempo, em face a hiperconectividade que os indivíduos estão expostos, revela o ritmo da atualidade que é acelerado ou hiperacelerado.

Na modernidade acreditava-se que o progresso da ciência proporcionaria um futuro onde seria alcançada a felicidade e as respostas para os anseios daquela época. No entanto, alguns desencaminhamentos históricos frustraram essa expectativa positiva do futuro, contribuindo para o fracasso da modernidade, como foi o caso da II Guerra Mundial. Este desapontamento com o futuro, ao avançar na chamada pós-modernidade, conferiu prestígio e valorização ao presente. Nesse momento, houve uma recusa em pensar sobre o que viria no futuro e, ao mesmo tempo, as práticas do passado passaram a ser negadas. A fase pós-moderna “indicava o advento de uma temporalidade social inédita, marcada pela primazia do anti-agora” (LIPOVETSKY, 2004, p.51).

Passado da pós-modernidade para a hipermodernidade, na década de 1990, o culto ao tempo presente é mantido, porém, as relações com as práticas do passado não são abandonadas e há interesse pelas questões do futuro. O passado para os indivíduos hipermodernos pressupõe um referencial de segurança, adaptado à lógica da moda – na qual “a antiguidade e a nostalgia se tornaram argumentos comerciais, ferramentas

mercadológicas” (LIPOVETSKY, 2004, p. 89). Já o futuro gera reflexões sobre incertezas e inseguranças por não ser possível prever o que está por vir.

O tempo, na hipermodernidade, é considerado escasso e comprimido, mas é nele que as pessoas têm de fazer valer o conceito da eficiência – realizar a maior quantidade de atividades com elevado nível de qualidade, no menor tempo possível. “Na hipermodernidade, não há escolha, não há alternativa, senão evoluir, acelerar para não ser ultrapassado pela ‘evolução’: o culto da modernização técnica prevaleceu sobre a glorificação dos fins e dos ideais.” (LIPOVETSKY, 2004, p.57).

Ao mesmo tempo que o novo é valorizado, o indivíduo hipermoderno resgata o antigo e está preocupado com o que está por vir. Mesmo com um tempo escasso e hiperacelerado, há brechas para buscar a qualidade de vida e cuidados com o corpo, como reflexo da apreensão com o futuro incerto. Esse ritmo de buscas indica que o regalo ao presente, antes vivido na pós-modernidade, deu lugar à angustia, na hipermodernidade. Trata-se de um período de “incerteza existencial, medo do futuro, uma perpétua ansiedade e uma sensação de urgência sem fim”. (BAUMAN, 2016).

Mesmo com a perturbação do cenário globalizado, Bauman (2005) diz que as pessoas seguem buscando continuamente suas identidades, pois são motivadas pelo desejo de segurança – corroborando o pensamento de Lipovetsky (2004). A segurança da modernidade que deu lugar à liberdade contemporânea, agora, faz falta. Isso porque o sujeito hipermoderno convive com paradoxos diários, pois são ao mesmo tempo “mais informados e mais desestruturados, mais adultos e mais instáveis, menos ideológicos e mais tributários das modas, mais abertos e mais influenciáveis, mais críticos e mais superficiais, mais céticos e menos profundos” (LIPOVETSKY, 2004, p.27-28).

O indivíduo se mostra cada vez mais aberto e cambiante, fluido e socialmente independente. Mas essa volatilidade significa muito mais a desestabilização do eu do que a afirmação triunfante de um indivíduo que é senhor de si mesmo. Testemunho disso é a maré montante de sintomas psicossomáticos, de distúrbios compulsivos, de depressões, de crescente sentimento de insuficiência e auto depreciação. (LIPOVETSKY, 2004, p. 83).

Embora permeado de contradições, o indivíduo hipermoderno vive na tentativa de aproveitar a vida em plenitude, de obter realização pessoal e imediata, o que converge com os valores da “sociedade de consumidores” de Baumaun (2008), na qual o sujeito

transforma-se em mercadoria vendável e anseia pelo seu melhor desempenho. Em prol da máxima aceitação, o sujeito mercadoria exige alternância de identidades conforme a demanda social, o que reforça sua contemporaneidade líquida e efêmera, de forma que se apresenta como resposta à necessidade ou desejo do outro.

Diferentemente de uma mercadoria, o sujeito mercadoria não se limita a expor como atributo de valor apenas aspectos físicos, Lipovetsky (2007) faz referência a uma etapa do consumo caracterizada pelo hiperconsumismo, o qual não comercializa produtos, mas estilo de vida.

Hoje o consumidor que era vulnerável às coerções sociais passou “a um hiperconsumidor à espreita de experiências emocionais e de maior bem-estar, de qualidade de vida e de saúde, de marcas e de autenticidade, de imediatismo e de comunicação.” (LIPOVETSKY, 2007, p. 14).

A dinâmica comercial na pós-modernidade admite que o indivíduo consuma quase tudo o que deseja, mas ele pouco percebe que também está sendo consumido pelo mercado orientado pela lógica capitalista. As empresas, de modo geral, tendem a traçar estratégias para conhecer qual o público que poderá gerar maior lucratividade e, nesse processo, lançam mecanismos que fazem do consumidor uma verdadeira mercadoria. Na obra “Vida para consumo” de Zygmund Bauman (2008), o sociólogo apresenta esse processo de transformação do consumidor em mercadoria e exemplifica algumas situações comuns do dia a dia em que as pessoas são

[...] aliciada, estimuladas ou forçadas a promover uma mercadoria atraente e desejável. Para tanto, fazem o máximo possível e usam os melhores recursos que tem a disposição para aumentar o valor de mercado dos produtos que estão vendendo. E os produtos que são encorajadas a colocar no mercado, promover e vender são elas mesmas. (BAUMAN, 2008, p. 13).

A essa organização do mercado, Bauman (2008, p.20) dá o nome de “sociedade dos consumidores” e considera que não há mais uma fronteira que separa “as coisas a serem consumidas e os seres humanos que as consomem”, pois todos somos “mercadorias vendáveis”. Bauman (2008, p.76) define esse processo em que o consumidor é elevado a condição de mercadoria como “comodização” e entende isso como sendo crucial no consumo, pois vai além da satisfação da necessidade, dos desejos e das vontades.

Neste cenário de hiperconectividade e da busca pelo seu melhor desempenho, valorizando o presente, buscando o prazer instantâneo, mas também preocupado com o futuro incerto, a internet, por meio das redes sociais propicia não só um ambiente de contato e troca de informações, como um espaço para a auto-exposição e hiperconsumismo, configurando-se como vitrine ideal, uma vez que permite a exposição de largo alcance das identidades a serem construídas e do estilo de vida a ser comercializado.

Não atoa verifica-se o aumento crescente no número de usuários das redes, bem como o convívio de diferentes redes sociais destinadas a finalidades específicas, como carreira e vida social. Reforçando esta perspectiva, a grande quantidade de tempo dispendido em redes sociais, mesmo em um período de sobrecarga de atividades, revela a sua importância para esta sociedade hiperconectada.

Ademais do ambiente ideal de exposição pessoal, Castells (1999) destaca a transformação social promovida pela formação de rede, indicando uma mudança na morfologia social do mundo contemporâneo. As redes possuem o poder de agregar indivíduos e promover o diálogo e a troca de valores e experiências e, a sociedade organizada em rede pode utilizar dessa estrutura para se solidarizar com o outro, ser canal de apoio e cooperação, ou pode ser utilizada para exclusão e controle. De acordo com Castells (1999), nesta nova estrutura social o poder reside na mente dos indivíduos. Nesta nova forma de organização, o poder se redistribui: cada nó da rede transforma-se em um centro de poder, ainda que difuso. O autor afirma:

O poder já não se concentra nas instituições (o Estado), nas organizações (empresas capitalistas) ou nos mecanismos simbólicos de controle (empresas midiáticas, igrejas). Pelo contrário, difunde-se em redes globais de riqueza, poder, informação e imagens, que circulam e se transmitem num sistema de geometria variável e geografia desmaterializada. No entanto, o poder não desaparece. O poder ainda governa a sociedade; ainda nos molda e domina (...) A nova forma de poder reside nos códigos da informação e nas imagens da representação em torno das quais as sociedades organizam as suas instituições e as pessoas constroem as suas vidas e decidem o seu comportamento. Esse poder encontra-se nas mentes das pessoas (CASTELLS, 1999, p.505-506).

Castells (1999) afirma que os sujeitos protagonistas das transformações sociais na era da informação são os movimentos sociais. Os movimentos, diz Castells (1999),

devem atuar através da rede, garantindo uma intervenção de forma descentralizada e integrada. Os líderes desses movimentos são meramente simbólicos, uma vez que o poder é a própria rede. As transformações da realidade física se dão a partir da atividade virtual que os indivíduos envolvidos articulam e produzem.

Desta forma, observa-se, como parte das transformações da sociedade, que os meios de comunicação ganharam maior espaço, passando a estar imbuído nela, permeando os relacionamentos e expressões. Este processo, tido como midiatização, se assemelha à globalização no que se refere a sua manifestação global e com influências visíveis na sociedade como um todo. Segundo Hjarvard (2014) a midiatização pode ser considerada um processo macrossocial, de forma que é necessário compreender seu funcionamento para melhor compreender a cultura e a sociedade.

Dada a complexidade que permeia o uso e práticas das mídias na atualidade, compreender sua lógica corrobora com o entendimento da sociedade que faz uso e a constitui, sendo as mídias sociais peça fundamental, uma vez que estão diretamente relacionadas com os indivíduos em suas relações sociais, seja nos vínculos afetivos, de trabalho e com as instituições.

## **2.2 MIDIATIZAÇÃO E MÍDIA SOCIAL**

Nas últimas décadas, grande parte das mudanças na sociedade se deu ao avanço tecnológico e à forma de comunicação mais acessível e rápida. Isso impactou diretamente na forma de pensar e agir das pessoas, principalmente, em seu comportamento social. O desenvolvimento de TICs advindas do uso da internet impactaram diversas áreas do conhecimento devido à ampliação de conexão em rede e às mudanças nas formas de se relacionar, se informar, se entreter, trabalhar e consumir.

Com o desenvolvimento das primeiras TICs, como rádio e televisão, o termo mídia começou a ser comumente empregado para designar todos os meios de comunicação. “De meados dos anos 90 para cá, a emergência da comunicação planetária via redes de teleinformática acabou por instalar a generalização do emprego da palavra “mídia” para se referir também a todos os processos de comunicação mediados por computador”. (SANTAELLA, 2003, p 54).

Inicialmente a mídia era concebida de forma dissociada da cultura e da sociedade, exercendo uma influência sobre ambas ou funcionando como um instrumento a ser

utilizado pelos indivíduos ou instituições para diferentes fins. Contudo, a sociedade contemporânea está permeada pela mídia, dificultando sua análise isolada. Neste contexto, a teoria da midiatização busca compreender de que forma as instituições sociais e os processos culturais mudaram de caráter, função e estrutura em resposta à onipresença dos meios de comunicação. (Hjarvard, 2014)

Tradicionalmente os estudos de mídia e comunicação tinham como objeto de pesquisa o processo de comunicação propriamente dito e, por isso eram abordados de forma independente. Contudo, em uma sociedade na qual os meios de comunicações estão tão presentes e entremeados com a cultura, o estudo das mídias é parte integrante de análise para melhor compreendê-la.

“Significa que, para compreender a importância da mídia em nossa moderna cultura e sociedade, já não podemos contar com modelos que a concebem isoladamente da cultura e da sociedade, ou que se limitam a considerar tão somente o processo de mediação.” Hjarvard, 2014, p15.

Pelo fato dos meios de comunicação estarem presentes em diferentes esferas sociais e culturais e, pelo processo de midiatização implicar em todas as áreas, seu estudo requer o olhar interdisciplinar.

“O contributo da teoria da midiatização para esse empreendimento interdisciplinar está em fornecer um arcabouço para analisar e construir uma compreensão teórica acerca das possibilidades de interações dos meios de comunicação com outros processos sociais e culturais, bem como um conjunto de hipóteses sobre os possíveis resultados da crescente presença de diversas mídias na cultura e na sociedade.” (Hjarvard, 2014, p18)

O termo “midiatização” tem sido empregado para compreender a influência exercida pela mídia em diversos âmbitos, como para pesquisa, uma vez que “desempenham um importante papel na produção e difusão do conhecimento e das interpretações da ciência” (Hjarvard, 2014, p.26). Isso ocorre a medida que descobertas ou conceitos científicos transitam da esfera acadêmica e passam a fazer parte de narrativas presentes nas diversas mídias, como é o caso de um filme no qual o conteúdo

lúdico e informacional se mistura usando-se de artifícios tecnológicos de imagem e som com os aspectos ficcionais.

Rödder e Schäfer (2010) alertam que midiática da ciência está limitada a algumas disciplinas, cientistas e fases de pesquisa. Por isso, é compreendida como uma instituição menos midiática em comparação com outros setores da sociedade. Portanto, dada a participação da mídia na sociedade contemporânea, iniciativas de aproximação da ciência e sociedade passam pela necessidade de maior interação midiática.

Conforme visão de James B. Thompson (1990, 1995), na qual a midiática é um elemento constitutivo do desenvolvimento da sociedade, as transformações observadas atualmente são fruto de um contínuo processo. A invenção da prensa tipográfica, na metade do século XV, revolucionou ao possibilitar o armazenamento, acúmulo e circulação de informação em uma escala sem precedentes e institucionalizou os meios de comunicação em massa, transformando o que antes era restrito ao encontro físico a uma comunicação mediada.

As inovações nos meios de comunicação corroboraram com as transformações na sociedade, que antes era agrária e feudal, e auxiliaram no processo e surgimento de instituições modernas. A televisão, mais do que o rádio, revolucionou a cultura, os comportamentos, a economia e conseqüentemente a publicidade. As mensagens, antes transmitidas utilizando apenas o som, passaram a ganhar imagens e movimentos, ampliando seu poder de transmitir informações e significados. Desta forma se impôs como um meio de comunicação hegemônico e capaz de possibilitar “uma uniformização dos padrões referenciais de consumo, nos quais praticamente as mesmas mercadorias seriam desejadas, independentemente do grupo ao qual o indivíduo pertença” (LYRA, 2001, p.4).

Ao longo do século XX o telefone, o cinema, o rádio, as revistas e a televisão compunham um sistema dominante de transmissão de ideias, valores e desejos de consumo. Ambos os meios de comunicação estabeleciam um relacionamento de transmissão de informações em massa, sem ou com baixíssima participação do receptor, que exercia um papel passivo nesta comunicação.

A partir da década de 1990, com a chegada da internet, é introduzido a mídia digital ou eletrônica que, entre outros fatores, confere maior controle da informação no

receptor por meio de sua participação na construção narrativa e da interatividade. Através do recurso do hipertexto, no qual por meio de links o emissor pode construir uma navegação única e individual, acarreta em uma nova concepção de aprendizagem e troca de informação a partir de sua teia não-linear e suas múltiplas possibilidades de interação. (Ferrari, 2010)

“O digital , ao contrário da comunicação escrita que se encerra no momento da impressão, potencializa o alcance e amplia as opções de leitura, permitindo que o leitor, ou usuário, assuma o papel de comando, reformulando as telas hipertextuais que contêm textos, imagens, simulações interativas e bases de dados em produto personalizado, sob medida.” (Ferrari, 2010, p.102)

Da mesma forma, o hipertexto científico e documental, com alto teor cumulativo, foi beneficiado com a chegada da tecnologia digital, pois a partir da indexação de conteúdos em estruturas de banco de dados, é possível armazenar ordenadamente conhecimento e cruzar dados.

Além de seu caráter hipertextual, há o fator coletivo e colaborativo nas mídias digitais. A participação do emissor e a individualização da comunicação é algo acelerado, fruto de um movimento que foi se intensificando à medida que novos recursos tecnológicos foram lançados. No início dos anos 80, começaram a surgir dispositivos que possibilitaram a individualização do sujeito, ou seja, aos poucos foram ganhando popularidade produtos que permitiam ao receptor escolher o que ouvir, assistir ou ler. Surgiram o walkman e as videolocadoras, que tornaram possível a escolha individualizada de músicas e filmes, culminando com o ápice da individualização com o surgimento da TV a cabo. (SANTAELLA, 2003)

Para Pierre Lévy (1999), o “ciberespaço permite, ao mesmo tempo, a reciprocidade da comunicação e a partilha de um contexto. Trata-se de comunicação conforme um dispositivo de todos para todos”. A participação do usuário e a personalização da mídia são possibilidades tecnológicas alinhadas com as transformações e anseios da sociedade que busca o consumo individualizado e não massivo, que era muito presente na sociedade fabril.

A mídia digital, com o lançamento de smartphones, em especial o iPhone em 2007, tem uma dimensão de interação aumentada. O aparelho ganhou larga aceitação e inaugurou um novo formato de comunicação, abrindo espaço para mais inovações através dos aplicativos.

Anterior ao desenvolvimento do computador e seus subsequentes, a máquina era vista com uma funcionalidade específica, utilizada quando necessário. Entretanto, o smartphone, agrega funcionalidades de outros aparelhos e novos recursos, tornando-se um item pessoal e indispensável, uma extensão do próprio corpo. Conforme Alves (2017, p.24) “A máquina, mais do que nunca, era uma extensão do nosso corpo; estaria para sempre acoplada a nós”.

Deste modo, o uso da internet e seus desdobramentos tecnológicos aproximaram a mídia do cotidiano, impactando o *modus operandi* das instituições, cultura e sociedade. Segundo Baudrillard (1994; 175f) as representações midiáticas da realidade adquiriram tal predomínio na nossa sociedade, que tanto as percepções e construções da realidade quanto o comportamento tem como partida representações mediadas, guiadas pelos meios de comunicação. Ocorrendo um processo de midiatização, quando os meios de comunicação exercem uma influência dominante sobre outras instituições.

Os sites de redes sociais se tornaram desde o início dos anos 2000 um fenômeno central da cibercultura, afetando as práticas socio comunicativas cotidianas em escala mundial. Elas se diferenciam dos demais meios de comunicação uma vez que nascem de um conceito descentralizado, ou seja, não há a priori um interesse central de um único emissor, são formas de organização humana e de articulação entre grupos e instituições que estão intimamente vinculadas ao avanço das redes físicas e dos recursos comunicativos. Ou seja, é um fenômeno que surge anterior aos recursos tecnológicos da internet e que é potencializado ao fazer uso das TICs.

Com o surgimento e disseminação das redes sociais, o conceito de comunidades, estudado por áreas como sociologia e antropologia, tradicionalmente definido como entidades geograficamente circunscritas, sofreu modificação uma vez que a rede de computadores proporcionou a criação de comunidades desconsiderando as barreiras geográficas e de tempo, mas levando em conta apenas as afinidades de interesses (KOZINETS, 1999).

As reuniões em grupos de interesse têm recebido diferentes denominações. Rheingold (1993) definiu-a como *comunidades virtuais*, uma agregação social em que ocorrem discussões públicas por um determinado tempo e com iterações humanas suficientes para criar conexões, redes de relacionamento pessoais no ciberespaço. Estas *comunidades virtuais* possuem diversos efeitos culturais, podendo influenciar nos hábitos e decisões, inclusive no de consumo.

Os usuários da internet participam de comunidades virtuais em busca de informações, amizade e apoio social, assim como o que ocorre sem o uso da internet, o que diferencia é a possibilidade de se acompanhar como se dá as relações em uma comunidade virtual, despertando o interesse de estudiosos e de organizações interessadas em melhor conhecer estes usuários (RIDINGS e GEFEN, 2004).

De forma similar, pode-se referir às comunidades virtuais como redes sociais, uma vez que são uma estrutura social composta por pessoas e/ou organizações, conectadas em rede por um ou vários tipos de relação, que partilham valores e interesses em comuns. O inicial objetivo de uma simples comunicação entre os membros ganhou características e formato específicos, observado pelo aumento na quantidade e particularidades das redes sociais, retratado na Tabela 1, que possibilitaram também que um mesmo indivíduo faça parte de diferentes redes e as utilize com diferentes propósitos como o profissional, social, afetivo, entre outros.

**Tabela 1. Principais mídias sociais do mundo**

<b>Rede Social</b>	<b>Ano de lançamento</b>	<b>País de origem</b>	<b>Número estimado de usuários</b>
Facebook	2004	Estados Unidos	2.20 bilhões de usuários ativos por mês
You Tube	2005	Estados Unidos	1 bilhão de usuários
Instagram	2010	Estados Unidos	800 milhões de usuários
Qzone	2005	China	600 milhões de usuários
Weibo	2009	China	600 milhões de usuários
Twitter	2006	Estados Unidos	330 milhões de usuários ativos por mês

Fonte: site das redes sociais.

Além dos portais das principais redes de rádio e TV, jornais e revistas tradicionais, as redes sociais como Facebook, Twitter entre outras, também permitem a rápida

disseminação de informações: “Por estarem conectadas, todas as informações que os usuários visualizam (notícias, novidades, vídeos, leituras, links, entrevistas, entre outros) podem ser rapidamente compartilhadas, o que vem a caracterizar a participação do sujeito na rede, criando sua identidade no ciberespaço” (MARCON; MACHADO; CARVALHO, 2012, p. 3).

Como desdobramento do uso do termo mídia, a cultura midiática, dentre algumas características, propicia o cruzamento de identidades uma vez que favorece a articulação em distintos níveis, de formas de culturas. Seu desenvolvimento está intrínseco ao avanço de novas tecnologias de comunicação e é responsável pelo aumento de mercados culturais e pela criação e expansão de novos hábitos de consumo de cultura.

Como se não bastassem as instabilidades, interstícios, deslizamentos e reorganizações constantes dos cenários culturais midiáticos pós-modernos, desde meados dos anos 90, esses cenários começaram a conviver com uma revolução da informação e da comunicação cada vez mais onipresente que vem sendo chamada de revolução digital. (Santaella, 2003, p 59)

As estratégias de persuasão na mídia voltada para o consumo têm se apresentado cada vez mais complexas. Influência social é definida como o processo em que os pensamentos, sentimentos e comportamentos de uma pessoa são afetados por uma outra pessoa ou por um grupo (Forgas & Williams, 2001). É um fenômeno transversal que se expressa por diversas formas, tais como a obediência, a conformidade, a atração interpessoal e a identificação. Uma dessas formas se configura como persuasão, quando um agente tenta mudar intencionalmente a atitude ou o comportamento de outras pessoas, sem que haja coesão ou qualquer relação de poder entre as partes (Perloff, 2010). Quando se analisa a publicidade voltada para promover um determinado consumo, portanto, deve-se referir a estratégias persuasivas, como objetivo final de uma série de táticas específicas de influência social.

Cialdini (1998) argumenta que grande parte da influência social acontece por meio de respostas automáticas de quem é influenciado e que as pistas que eliciam essas respostas podem ser identificadas em diferentes culturas quando se tem o intuito de estimular pessoas a atender ou concordar com um pedido.

Para o filósofo, a motivação para tornar a inteligência coletiva mais ampla, passa pela prosperidade econômica, educacional e social, fortalecendo os direitos humanos e

a transmissão das heranças culturais – gerando um planeta mais sustentável. Hoje a vida humana passa pelo digital. Não dá, por exemplo, para tratar dos dentes num dentista virtual, ou cortar o cabelo pela web, mas tudo será armazenado na internet e a computação baseada em TAG vai modificar nosso dia a dia, criando inteligência coletiva reflexiva, ou melhor, autorreflexiva, nos ensina Levy.

### **2.3 BIG DATA, MUITO ALÉM DE BANCO DE DADOS**

As relações de trabalho, lazer, educação e a sociedade como um todo podem refletir em dados da mesma forma com também são afetados pelo seu volume, variedade e, principalmente, velocidade com que crescem e trafegam, principalmente quando em meio digital. Esta realidade impacta todos os setores econômicos, desde um simples consultório médico até grandes indústrias varejistas. Isso se dá pois, no ciberespaço as informações são geradas de maneira contínua e em todos os tipos de formatos: texto, imagem, vídeo, entre outras, ocasionando um aumento progressivo do volume de informações produzidas diariamente.

Dado estas proporções torna-se urgente o desenvolvimento de padrões de organização de forma a disponibilizar a informação armazenada e o aperfeiçoamento das técnicas utilizadas para análise e administração dos dados (CALDAS e SILVA, 2016). Atrelado ao conceito de organização dos dados há também a demanda por “*datafication*”, ou seja, transformar a informação em dado quantificável de forma a permitir sua análise e cruzamento com outros dados.

De acordo com Lima Junior (2011), a expressão Big Data se refere ao conjunto de dados produzidos atualmente no mundo, cujo volume está além dos padrões e da capacidade das ferramentas utilizadas por bancos de dados para capturá-los, analisá-los e gerenciá-los. O Big Data relaciona os dados de forma a apresentar uma interação entre eles e não uma análise de sua causa. Sendo assim, o tratamento desses dados auxilia a pontar a ocorrência estatística de um fato fornecendo insumos para projeções, descobertas e plano ações de forma mais estratégica. (CUKIER & MAYER-SCHOENBERGER, 2013)

O volume de dados está crescendo de forma exponencial e pode chegar a 16 zetabytes (16 trilhões de GB) em 2020 (Turner, 2014) e a utilização do Big Data

acompanha o mesmo ritmo acelerado, firmando-se como instrumento fundamental em todas as áreas. Boyd e Crawford (2012) consideram que o grande volume de informações disponíveis na rede é capaz de alterar a própria constituição do conhecimento. Não se trata apenas de grandes conjuntos de dados, mas uma mudança no pensamento que se assemelha a transformação provocada por Ford com relação à forma de produção e organização do trabalho.

Schönberger-Mayer e Cukier (2013) complementam com a constatação de que o volume de informação cresceu tanto que passou a não caber na memória de processamento dos computadores, por isso os engenheiros tiveram de aprimorar os instrumentos utilizados para fazer a análise. Esta é a origem de novas tecnologias de processamento como a MapReduce da Google e sua equivalente de código aberto, Hadoop, lançada pela Yahoo.

Os dados começaram a se acumular aproximadamente meio século depois de os computadores se difundirem no meio social. A Tabela 02 mostra a evolução da utilização dos dados e informação desde sua origem até os dias atuais:

**Tabela 2. Evolução da utilização dos dados e informação**

<b>Ano ou período</b>	<b>Fato</b>
8000 a.C.	Comerciantes sumérios usavam tábuas e fichas de barro para denotar os bens comercializados.
1086	Livro Domesday, um dos mais venerados tesouros britânicos, foi uma contagem abrangente - e sem precedentes - dos ingleses, suas terras e propriedades.
1439	Primeira revolução da informação: Impressora de Gutenberg.
1453 a 1503	De acordo com a historiadora Elizabeth Eisenstein, oito milhões de livros foram impressos. Volume de informações dobra na Europa, em 50 anos.
1662	O comerciante britânico chamado John Graunt queria saber a população de Londres na época da peste negra. Em vez de contar as pessoas, ele inventou uma abordagem - que hoje chamamos de "estatística"- que lhe permitiu estimar a população.

1880	O Census Bureau dos Estados Unidos contratou o inventor Herman Hollerith, que desenvolveu um sistema de cartões perfurados e máquinas de tabulação para o censo de 1890. Ele conseguiu diminuir o tempo de tabulação do censo de oito anos para um.*
1920	Descobertas da mecânica quântica abalaram para sempre o sonho das medições abrangentes e perfeitas.
1934	Jerzy Neyman, estatístico polonês, demonstrou que a abordagem da amostragem levava a erros e que para isto deveria ser considerada a aleatoriedade na escolha da amostra. A amostragem tornou-se a solução para o problema da profusão de dados.**
1950-1960	Um sistema de folha de pagamento foi implementado na General Electric Corporation, por Joe Glickauf e Arthur Andersen em um computador eletrônico digital. Início das ideias de "Revolução da informação" e "Era Digital" ou "Era da informação".
1980-1990	Nesta época, Leonard Kleinrock, Lawrence Roberts, Robert Kahn, and Vint Cerf inventaram a internet. 40% de toda a potência computacional do mundo existia na forma de calculadoras de bolso. "Era da rede".
2000	Dados: 75% dados analógicos (papel, filme, vinil, fitas magnéticas, livros, fotografia); 25% dados digitais.
2007	Dados: 7% dados analógicos; 93% dados digitais.
2010	Ciências como a astronomia e a genômica, vivenciaram uma explosão informacional e cunharam o termo big data. Este fenômeno, representa uma transição em termos de armazenamento e análise. "Era big data".
2013	Dados: 2% dados analógicos; 98% dados digitais - 1200 exabytes. Volume de informações dobra a cada 3 anos.

Fonte: Nesello,P com base em Schönberger-Mayer e Cukier (2013) e Mineli, Chambers e Dhiraj (2013). \* Marco do início do processamento automatizado de dados: a quantidade de informações coletadas emperrava os instrumentos usados para processá-la, e eram necessárias novas técnicas. \*\* A amostragem aleatório foi um sucesso e é a base da medição moderna em escala.

Sendo assim, através da Tabela 2 pode-se verificar que há uma grande quantidade de informação disponível, ao mesmo tempo que ela se acumula com maior rapidez, além disso, verifica-se que os avanços tecnológicos em TI que possibilitaram o big data é fruto de uma contínua busca da humanidade por medir, registrar e analisar o mundo (SCHÖNBERGER-MAYER E CUKIER, 2013). Diferentemente do interesse de compreender a causalidade, segundo proposto por Schönberger-Mayer e Cukier (2013), o big data objetiva extrair novas ideias e criar novas formas de valor de maneira que se alterem os mercados através da análise de grande escala, que não poderiam ser feitos em escala menor.

A princípio, pode-se definir o conceito de Big Data como sendo um conjunto de dados extremamente amplos e que, por este motivo, necessitam de ferramentas especialmente preparadas para lidar com grandes volumes, de forma que toda e qualquer informação processada por esses meios possa ser encontrada, analisada e aproveitada em tempo hábil: “O valor real do Big Data está no insight que ele produz quando analisado – buscando padrões, derivando significado, tomando decisões e, por fim, respondendo ao mundo com inteligência” (INTEL, 2013, p. 3).

Resgatando os precursores mais recentes, uma década antes do surgimento do Big Data verifica-se inicialmente o conceito de Data Warehouse em 1980. A diferença de Data Warehouse e Big Data é a velocidade com que os dados precisam ser disponibilizados, uma vez que, em um projeto de Data Warehouse, o processo torna-se mais lento até que as informações estejam disponíveis. Segundo Taurion (2013), o Big Data está muito além de um grande Data Warehouse. Entretanto, pode-se dizer que o Big Data se relaciona com o Data Warehouse. Já na década de 1980, o volume de dados crescia exponencialmente. Foi o momento em que surgiram dúvidas de o que fazer com os dados armazenados. Na época, as técnicas utilizadas já não supriam as necessidades de tratar a maioria dos repositórios de dados. Com essa demanda crescente, no fim da década de 1980, surge o conceito e a utilização de Data Mining ou Mineração de Dados.

O volume crescente de dados a serem tratado coincide com o avanço das redes sociais, onde a cada minuto os servidores dessas plataformas são bombardeados com milhões de dados, além do Enterprise Resource Planning (Planejamento de Recursos Empresariais - ERP), que concentra dados de toda uma organização.

Outro diferencial dos recursos do Big Data é sua capacidade de analisar todos os tipos de dado, tanto dados estruturados (de um sistema ERP, por exemplo) como não

estruturados, extraídos principalmente, das redes sociais. Ao analisar esses dados não estruturados, que antes eram considerados sem valor, o Big Data expande a capacidade de compreensão dos acontecimentos, abrindo espaço para novas reflexões e decisões. Em vez de analisar apenas um percentual de dados, como uma amostragem, por exemplo, seriam analisados além da amostragem, dados de diversas fontes nunca utilizadas antes, como um comentário numa rede social. Um exemplo prático disso, são as pesquisas de boca de urna, em que os dados são analisados baseados em levantamentos de amostragem – se a apuração dos eleitores fosse completa, teoricamente, seria feita a eleição de fato.

Mais do que a quantidade de dados, conforme o termo Big Data sugere, a expressão é bem mais abrangente e tem como base 5 pilares, também conhecido com os 5 Vês: volume, velocidade, variedade, veracidade, valor. De forma simplificada estes pilares representam a grande quantidade de dados que são registrados a partir de simples transações cotidianas, a relação de velocidade que estes dados são produzidos e gravados, as diversas formas e formatos provenientes da possibilidade de se registrar e se transformar em dados textos, imagens e sons entre outros. Por sua vez a quantidade de dados e a possibilidade de produção deste fontes descentralizadas conferem também a preocupação pela confiabilidade da informação gerada, o que culminará e determinará o real valor destes dados transformando-se em informação.

O crescimento do processamento do Big Data vem se tornando amplo e complexo. Hoje existem infinitas opções e maneiras para uso de Big Data. Dados que podem afetar de forma positiva e/ou negativa todos os setores da economia. Os dados gerados são muitos e diversos, coletados de todo lugar, de todos os setores, econômico, educacional, lazer, saúde, etc. De um simples sensor num viaduto, uma compra no supermercado, uma ida ao médico para uma consulta rotineira, tudo gera dados, que constituem a matéria-prima de todas as análises e interpretações que podem ser geradas. As informações extraídas de um Big Data, geralmente, são apresentadas em aplicações destinadas aos usuários, com interface amigável e de fácil entendimento.

À medida que os dados são gerados e permitem produzir informações úteis, pode-se dizer que se está utilizando Big Data. Hoje ele já inserido em todos os setores econômicos e, a cada dia que passa, afeta ainda mais cada envolvido. Exemplos desse envolvimento é a participação mais ativa de tecnologias de Big Data no setor de saúde, educação e, principalmente, nas redes de varejistas e no mundo dos negócios. A cada

dia, empresas inovadoras como a IBM, LinkedIn, Amazon, entre outros, vêm desenvolvendo tecnologias e formas de utilização do Big Data, para as mais diversas áreas. As organizações cujo o produto fim não é a tecnologia, a usam cada vez mais, para obter diferenciação e se destacar das concorrentes.

O Big Data se expande para todos os campos do conhecimento humano. Apesar de seu conceito não ser tão claro, como já abordado, seu crescimento é volumoso, decorrente do enorme número de informações não estruturadas, além de ser necessário processamento em grande escala para que seja possível a extração de informações que possam estabelecer valores que impactam na economia, nas organizações, no governo e nas relações interpessoais.

Dentre as várias vantagens apontadas pelo uso de Big Data destaca-se a oportunidade de redução de custos, agilidade no processamento de tarefas e até mesmo a capacidade tornar possível a execução de tarefas e análises sem o qual não seria factível devido a quantidade de tempo e recursos que demandariam ser utilizados. Desta forma, o uso do Big Data oferece maior respaldo para a tomada de decisão, seja em qual for o setor que está endereçando. (Licks, 2016). E torna-se, portanto, base fundamental para o desenvolvimento de inovações, que por si carregam o potencial de desenvolvimento econômico e melhores condições de vida, além de benefícios ao meio ambiente como um todo.

Devido ao desafio da utilização inteligente dos dados, seja pelo seu volume crescente como também por seu potencial de utilização, o Big Data é um fenômeno de interesse de estudo em todas as áreas do conhecimento e inclusive auxilia na conexão entre eles, uma vez que cruza e analisa diferentes informações.

Na gestão de cidades, por exemplo, há a preocupação de torna-las mais competitivas, humanas e sustentáveis, aptas a resolver diversos problemas urbanos complexos como trânsito, segurança pública, poluição dentre outros, tendo como resultado maior a ser alcançado melhorar a qualidade de vida da população. As diversas características e habilidades que uma cidade deve possuir para melhorar seu desempenho em determinadas áreas, principalmente no gerenciamento de seus serviços, foi dado o nome Smart Cities, que se utiliza do Big Data como instrumento para melhor gestão de seus recursos e como fonte de informação para tomada de decisões mais assertivas. (GIFFINGER et al., 2007; VILAJOSANA et al., 2013; NEIROTTI et al., 2014)

As aplicações do uso do Big Data também encontram oportunidades na área da saúde pública. Conforme aponta Chiavegatto Filho (2015), a epidemiologia tem como desafio a análise de grandes bancos de dados (*big data*). Devido ao crescimento do número de estudos multicêntricos e, também a pressão pelo detalhamento dos gastos públicos, vê-se aumentada a quantidade de dados disponíveis e uma demanda por análise destes dados complexos e desestruturados. Desta forma o uso de recursos com maior capacidade de análise de grande quantidade de dados poderá auxiliar tanto na prevenção, identificação e melhor alocação dos recursos para oferecimento de melhor qualidade de saúde à população.

Por esses e outros exemplos, é possível identificar de forma crescente o uso e o valor do Big Data, cujo impacto tem despertado o interesse e, entre alguns estudos, Cavanillas, Curry e Wahlster (2016) estimam que as soluções utilizando Big Data podem contribuir para o aumento da competitividade europeia através de ferramentas, aplicações e serviços que agregam valor, com impacto direto na economia. Os autores apontam os principais empregos e impactos do Big Data nos diferentes setores:

Na saúde estudos utilizando Big Data podem contribuir para redução dos gastos na saúde pública através da identificação de incidência de doenças mais frequentes e suas correlações, direcionando ações preventivas ou fornecendo os insumos para medidas mais efetivas. Os benefícios no setor público incluem também o aumento da transparência através de dados abertos, melhorando a alocação de recursos, aumento da qualidade dos serviços e munindo a população com mais informações para exigir seus direitos.

No mercado financeiro e de seguros pode ser utilizado para análises preditivas e ser direcionado para ações de fidelização dos clientes e aumento de lucros, da mesma forma que pode ser empregado no varejo, identificando padrões de consumo e interesses relacionados. Além de auxiliar no processo de vendas de mercadorias e serviços, o Big Data também pode fornecer informações auxiliando no aumento da eficiência na produção, contribuindo para fábricas mais inteligentes e que desperdiçam menos recursos.

Dando continuidade aos destaques apontados por Cavanillas, Curry e Wahlster (2016), os benefícios abrangem também os setores de transportes. Uma vez que as necessidades de deslocamento são identificadas, todas as ações podem ser direcionadas de forma mais precisa, atendendo às necessidades da população. Na

mídia, telecomunicação e entretenimento vê no uso do Big Data a oportunidade de identificar gostos e direcionar suas ações, além de poder oferecer conteúdo mais personalizado, o que aumenta a relevância de sua oferta.

Devido ao seu potencial de ordenação e extração de valor de grande base de dados, emprego do Big Data realmente se aplica a todos os campos e está cada vez mais comum e necessário. Todos os tipos de dados que são monitorados hoje, e principalmente aqueles que se tornam possíveis de serem coletados e armazenados devido às evoluções tecnológicas, quando reunidos conferem um potencial informacional muito grande. Uma vez que a análise dos dados se incorpora nas práticas para a tomada de decisões, torna-se eminente e urgente a conversão dos diversos setores para o emprego do Big Data.

### 3. METODOLOGIA

Para a pesquisa utilizou-se abordagem qualitativa-quantitativa de caráter exploratório, adotando-se procedimentos bibliométricos complementados por análise de conteúdo, conforme metodologia apontada por Pereira (2009). Na perspectiva qualitativa o pesquisador utiliza-se de técnicas interpretativas, decodificando, ou seja, traduzindo o significado dos fenômenos observados como forma de análise, segundo assegura La Ville e Dionne (1999, p. 227): “o pesquisador decide prender-se às nuances de sentido que existem entre as unidades, aos elos lógicos entre essas unidades ou entre as categorias que reúnem”. Desta forma, a análise utiliza-se de inferências subjetivas, embora haja o esforço de um relato objetivado, explicado e justificado em cada etapa percorrida.

Por outro lado, a utilização da abordagem quantitativa possui características objetivas, havendo certo distanciamento necessário do pesquisador e de seus valores em face do objeto, como alerta Durkheim (1999). Além disso, a quantificação pode ser útil ao subsidiar o pesquisador com dados e estatísticas. Mesmo a quantificação sobre influências da complexa e às vezes inexplicável multifacetada realidade e, caso venha a obscurecer aspectos menos evidentes que permeiam as relações sociais, é reservada certa subjetividade na análise dos números revelados.

Sendo assim, por serem consideradas abordagens complementares, optou-se pela utilização de ambas as abordagens. Foram analisados todos os elementos condicionantes aos registros do conhecimento representados pelos artigos científicos de maneira qualitativa e quantitativa, caracterizando a pesquisa como descritiva, dado que para Gil (2002), as pesquisas descritivas têm como foco a descrição das características de determinada população ou fenômeno, permitindo, entre outras, o estabelecimento de relações entre variáveis.

Foram adotados procedimentos bibliométricos complementados por análise de conteúdo. Os procedimentos adotados estão relacionados aos Estudos Métricos da Informação no âmbito da coleta e análise de dados, que foram constituídos de artigos científicos produzidos sobre a temática. Em essência, os estudos métricos da informação destacam-se pela aplicação metodológica das técnicas bibliométricas e cientométricas utilizadas na pesquisa básica, na orientação científica de determinada especialidade, bem como na avaliação das estruturas da ciência (GLANZEL, 2003), a partir do registro de informações. O emprego desses estudos métricos auxiliou na compreensão da

ciência praticada, sua estrutura cognitiva e social (WHITE; McCAIN, 1998; CHEN, 2003; CHEN, 2001; GARFIELD, 1994; GARFIELD, 1964; PETTIGREW; McKECHNIE, 2001; PRICE, 1963; PRICE, 1965; PRICE, 1969; SMALL, 1977; WHITE, 1981; MOYA-ANEGON et al., 2006; MOYA-ANEGON et al., 2007; VARGAS-QUESADA et al., 2008; VARGAS-QUESADA et al., 2010).

Por meio do uso desses estudos, foi possível distinguir os trabalhos que centram esforços nas tendências científicas e tecnológicas emergentes (ASTROM, 2007; GALFIELD, 1994; CHEN, 2006; MORRIS; YEN; ASNAKE, 2003; PERSSON, 1994) e trabalhos que avaliam as mudanças bruscas do conhecimento científico (WHITE; McCAIN, 1998; CHEN, 2006; CHEN, 2003; CHEN, 2002; SMALL, 1980). Complementarmente, estudos desenvolvidos por pesquisadores brasileiros mostram que é possível compreender as relações entre produção científica e tecnológica (HYODO, 2010; MARICATO, 2010), traçar panorama dos estágios de desenvolvimento de determinados setores científicos, tecnológicos e industriais (PEREIRA; BAZI, 2009) de países e demais configurações sociais, bem como identificar estágios de institucionalização de determinada ciência (PEREIRA e FUJINO, 2014).

Foram realizadas leituras e análise de todos os registros informacionais. Para tanto, foi utilizado a técnica de “Análise de Conteúdo”, seguindo as propostas de Bardin (1977), a análise de conteúdo pode ser definida como um conjunto de técnicas de análise das comunicações visando obter, por procedimentos sistemáticos e objetivos de descrição do conteúdo das mensagens, indicadores de permitam a inferência de conhecimentos relativos às condições de produção e recepção destas mensagens, sendo um método de tratamento da informação contida nas mensagens, nesta pesquisa, os títulos e resumos dos artigos relacionados à temática.

Conforme Bardin (1977) o uso do método de análise de conteúdo permite ao pesquisador ultrapassar as incertezas e o enriquecimento da leitura, sem deixar de lado o rigor, a necessidade de descobrir, pelo questionamento. Ela tem por finalidade, a partir de um conjunto de técnicas parciais, mas complementares, explicar e sistematizar o conteúdo da mensagem e o significado desse conteúdo, por meio de deduções lógicas e justificadas. (OLIVEIRA et al., 2003).

A proposição da análise de conteúdo é a inferência de conhecimentos relativos às condições de produção ou de recepção (OLIVEIRA et al., 2003). Por Inferência, Bardin (1977) assinala que é a operação lógica pela qual se admite uma proposição em virtude

de uma ligação com outras proposições já aceitas como verdadeiras. Para a utilização do método foi necessária a criação de categorias relacionadas ao objeto de pesquisa, que, com base nelas, fez-se inferências, ou seja, foi possível extrair uma consequência, sendo deduzido de maneira lógica conhecimentos, o que permitindo a identificação das questões relevantes contidas no conteúdo das mensagens (OLIVEIRA et al., 2003).

Trabalhou-se em palavras que, isoladas, puderam atribuir relações com a mensagem ou possibilitar que se fosse feita inferência de conhecimento a partir da mensagem. Dessa maneira, estabeleceu-se correspondências entre as estruturas linguísticas ou semânticas e as estruturas psicológicas ou sociológicas de enunciados. Uma vez que, segundo Bardin (1977), a leitura do pesquisador responsável pela análise não é portanto, uma leitura à letra, mas, o realçar de um sentido que se encontra em segundo plano.

### **3.1. ESTRATÉGIA DE COLETA DOS DADOS**

Os trabalhos científicos publicados sobre *Big Data* e *mídias sociais* em diferentes campos do conhecimento formaram o *corpus* de análise do presente estudo. Para tal optou-se pela identificação dos artigos publicados em revistas indexadas na base de dados *Web of Science* (*Thomson Scientific / Institute for Scientific Information – ISI*).

A escolha dessa base de dados se justifica por ser uma base referencial multidisciplinar que cobre aproximadamente 12.000 periódicos científicos avaliados por critérios rigorosos a cada duas semanas. Além disso, é assegurado que os títulos mantenham altos padrões e uma clara relevância à comunidade científica por meio da constante revisão de cobertura de títulos existentes. Cada revista é submetida a um processo amplo de avaliação, por profissionais do corpo editorial da ISI com apropriada formação educacional, antes de ser selecionada ou rejeitada. No seu processo de avaliação, são considerados muitos fatores qualitativos e quantitativos, tais como: os padrões básicos de apresentação, o conteúdo editorial, a internacionalidade dos autores e a citação de dados associada a esses autores, para avaliação das revistas. Nenhum fator é considerado de forma isolado, mas, ao contrário, os dados são combinados e inter-relacionados, possibilitando ao editor a capacidade de determinar os pontos fortes e fracos de uma revista em sua totalidade (TESTA, 1998).

Dado o rigor da avaliação das revistas indexadas, a base de dados *Web of Science* é reconhecida como uma fonte que disponibiliza alta qualidade de informação no campo científico e tecnológico. Sendo assim, a base mostrou-se a mais adequada para esta pesquisa devido a amplitude dos campos científicos cobertos e pela qualidade das informações.

Para a busca dos artigos publicados em periódicos indexados na base de dados *Web of Science*, foi considerado tópico geral da pesquisa traduzido para o idioma inglês: “*Big Data e mídias sociais*”, realizando consulta conforme etapas indicadas por Lopes (2002).

Foram utilizadas palavras-chave por lógica booleana (palavras combinadas por AND e OR) e/ou por operadores de truncagem (palavras combinadas por ? ou \*, utilizados para recuperar palavras a partir dos prefixos ou sufixos ou recuperar palavras que apresentem grafias semelhantes). Todos os artigos publicados nos periódicos indexados pela base de dados *Web of Science* na data de 18 de abril de 2018 foram considerados, independentemente dos anos de publicação tomando-se como limite os artigos publicados até o ano de 2017.

Os 541 artigos recuperados formaram o *corpus*, e foram utilizados para a compreensão do estado da arte, dos comportamentos teórico e empírico identificados a partir dos critérios estabelecidos e submetidos a avaliação a partir dos critérios de verificação: *pergunta/problema/objetivos da pesquisa, aplicação/análise proposta, tipo de avaliação e resultados alcançados*.

Utilizou-se três estratégias de análise e interpretação qualitativa, enumeradas por La Ville e Dionne (1999):

- *Emparelhamento*: associação de dados recolhidos (registrados em um quadro operacional) discutido com a finalidade de compará-los;
- *Análise histórica*: elaborou-se um roteiro sobre a evolução do fenômeno ou da situação;
- *Construção interativa de uma explicação*: elaborou-se pouco a pouco uma explicação lógica do fenômeno ou da situação estudados, pois não supõe a pré-existência de um ponto de vista teórico.

### 3.2. PROCEDIMENTOS PARA ANÁLISE

Para a análise quantitativa dos dados representados nos registros informacionais foram utilizados softwares bibliométricos e cientométricos gratuitos, que se destinam ao processamento de informação científica e tecnológica e, permitem processar diversos documentos com técnicas bibliométricas avançadas, visualizar dados de diversas formas, agrupando-os em países, autores, instituições, temáticas, entre outros.

Desta forma, são ferramentas de tratamento dos dados, num processo conhecido como mineração de dados, e foram agrupados conforme interesse de análise. Segundo Hayashi (2004), é uma ferramenta de mineração de texto usada para transformar informação em conhecimento extraído de bases de dados. Trabalhos duplicados devido a indexação repetida de periódicos foram retirados e padronizados os campos de cada registro bibliográfico, a saber: nomes de autores, instituições, palavras-chave e referências. A padronização dos campos por meio do software foi de extrema importância, pois garantiu a qualidade do conjunto de dados.

Os dados e análises foram graficamente apresentados utilizando os programas *Microsoft Excel* e *Ucinet / Netdrawn*, além de gráficos fornecidos por softwares bibliométricos. Esses softwares são usualmente empregados na representação de indicadores relacionados às métricas de informação.

Foram realizadas leituras e análise de conteúdo de todos os títulos e resumos dos artigos, que permitiu a segmentação dos trabalhos em 3 diferentes categorias conforme segue:

- *Aspectos metodológicos*: artigos que apresentam modelos para análises ou que elucidam técnicas a serem utilizadas, bem como trabalhos que comparam diferentes metodologias;
- *Questionamentos e reflexões*: artigos que trazem uma análise social ou consequências do uso de big data em mídias sociais;
- *Estudos de casos*: artigos que se referem a alguma aplicação específica, restrita a uma situação em análise.

Após segmentação, os artigos foram classificados conforme os critérios estabelecidos e analisados.

#### 4. RESULTADOS E DISCUSSÃO

Realizou-se, conforme disposto na metodologia, busca dos artigos publicados em periódicos indexados na base de dados *Web of Science* cujo ano de publicação era igual ou anterior a 2017. Por meio da ferramenta de pesquisa disponível na base de dados, utilizou-se como filtro para a análise traduzido para o idioma inglês das palavras chaves: “*Big Data*” e “*Mídias Sociais*”.

Para melhor compreensão da temática, primeiramente buscou-se dimensionar a relevância do uso dos termos de forma isolada, ou seja analisou-se os artigos que empregam as palavras chaves em sua totalidade e, posteriormente, através do uso de lógica booleana (palavras combinadas por AND e OR), foi possível fazer o levantamento e análise dos artigos publicados nos periódicos indexados pela base de dados *Web of Science* que possuem ambos os termos com palavras chaves, demarcando portanto a temática foco deste estudo.

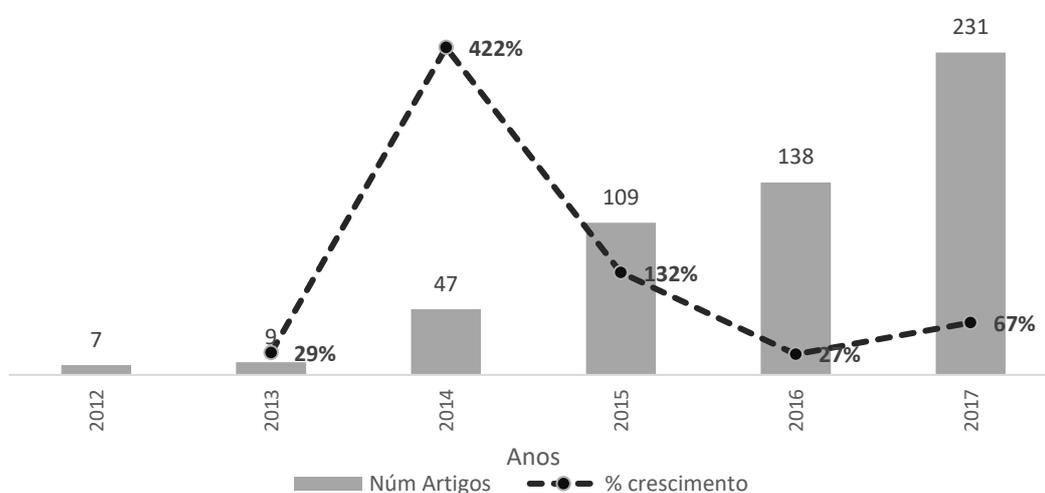
Verificou-se que ambos os termos começaram a aparecer, a partir da década de 90, como palavras chaves de artigos publicados. No mesmo período há o surgimento e popularização do uso de computadores pessoais e internet, o que pode sugerir uma relação entre os temas abordados nos estudos e o emprego desta tecnologia nova no cotidiano social, visto que a sua utilização poderia fazer emergir reflexões acerca de seus aspectos técnicos e, também sua relação com a sociedade.

Com relação à quantidade de artigos associados a estas palavras-chaves, o termo *Mídia Social* se mostrou mais frequente. Foram encontrados 15.277 artigos publicados até a data de 2017, enquanto que para o termo *Big Data*, até o mesmo ano, foi possível localizar 8.699 artigos, sendo, portanto, o correspondente a pouco mais da metade do primeiro (56,9%).

Por sua vez, ao analisar a temática em estudo, que considera os artigos que possuem ambos os termos como palavras-chave, identificou-se a publicação de 541 artigos. Com relação ao comportamento temporal da temática, verificou-se que o primeiro artigo foi publicado em 2012, quase 20 anos depois da primeira publicação de um artigo considerando apenas um dos termos como palavra-chave, que ocorreu em 1993 relacionado a *Big Data*. Sendo assim, verificou-se, que a temática é bastante recente e representa uma fatia ainda diminuta, 2%, quando comparada a todo o arcabouço identificado considerando o uso dos termos isoladamente (23.976 artigos que possuem *Big Data* ou *Mídias Sociais* como palavras-chave).

Por outro lado, verificou-se uma tendência de crescimento acelerada no número de artigos e aumento na participação da temática. Considerando apenas os trabalhos publicados no ano de 2017 observa-se que a temática representa 3% do total de artigos que relacionam algum dos termos.

Este aumento foi fruto do contínuo crescimento no número de publicações sobre a temática, verificado pela representatividade da produção do último ano (o número de artigos publicados em 2017 representa mais de 40% do total) e as taxas de crescimento anuais, que embora com alguma oscilação foram sempre superiores a 25%, conforme representado no gráfico a seguir:

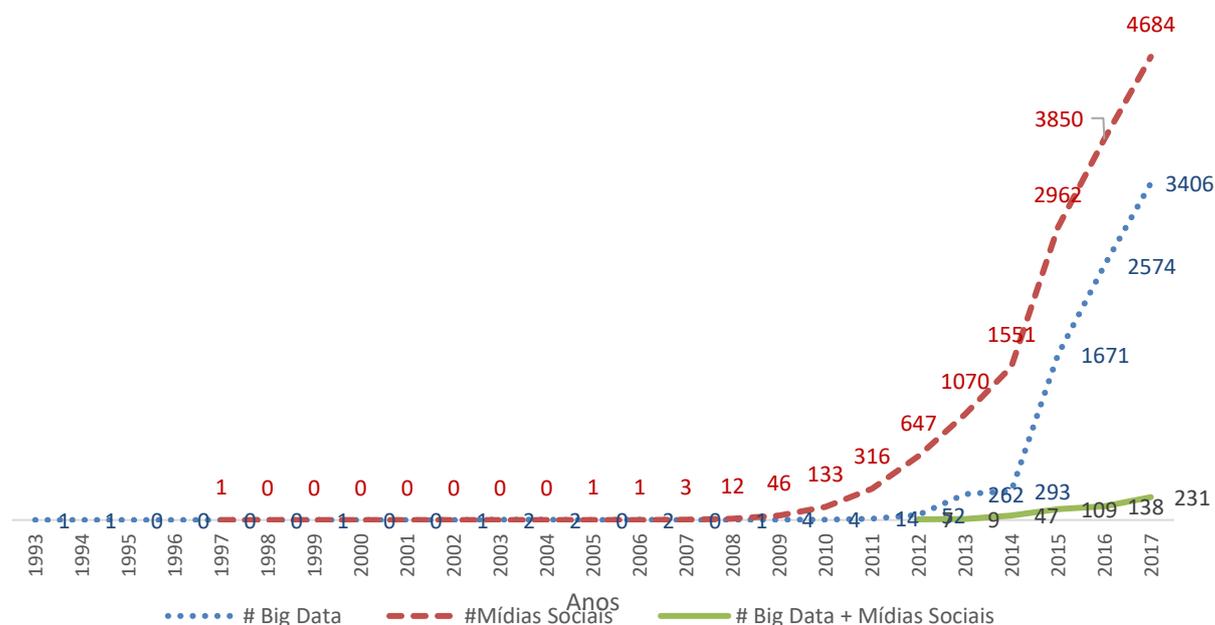


**Gráfico 1.** Número de artigos por ano e taxa de crescimento

Pode-se constatar que o número de artigos sobre a temática vem crescendo consistentemente desde 2012, tendo o pico de crescimento em 2014 (422% em relação ao ano anterior). Após este ano a taxa de crescimento no número de artigos sobre a temática desacelerou um pouco e voltou a crescer no último ano. Embora a taxa de crescimento tenha variado, verifica-se um crescimento no número de artigos desde o início das primeiras publicações sobre a temática, sugerindo um crescimento exponencial.

Quando sobrepostos a evolução de trabalhos publicados considerando os termos isoladamente verificou-se a tendência ainda mais acelerada. A partir de 2010 para *Mídias Sociais* e 2014 para *Big Data* a publicação de artigos iniciam um comportamento exponencial, ou seja, o aumento no número de artigos que até então apresentava um crescimento mais lento, passou a ganhar mais expressividade. Este comportamento

revelou que a temática está ganhando relevância e não se configura como um assunto de pontual e passageiro. O aumento considerável no número de artigos abordando os temas isoladamente ajudam a melhor compreender o crescimento da temática, conforme pode ser observado no gráfico a seguir.



**Gráfico 2.** Número de artigos por ano sobrepostos Big Data, Mídias Sociais e sua combinação

Pelo gráfico visualizou-se que o comportamento exponencial do termo mídia social iniciou antes que o mesmo ocorresse com o termo Big Data, fazendo com que, embora fosse mais recente, o número total de artigos seja maior. Este último, por sua vez, mostrou uma evolução bastante significativa a partir de 2014, quando saltou de 293 artigos para 1671.

Para melhor compreender de que forma o conhecimento científico sobre a temática vem sendo comunicado verificou-se os meios de sua publicação. Os 541 artigos analisados foram publicados em 344 periódicos distintos, o que confere uma média de 1,57 artigos por periódico. Este número indica uma pulverização da temática que se confirma ao verificar que 72% dos 344 periódicos publicaram apenas um único artigo. Considerando os periódicos que tiveram 5 ou mais artigos publicados sobre a temática, observou-se que se tratam de 13 periódicos e que foram responsáveis pela publicação de 89 artigos, ou seja pouco mais de 16% do total.

O periódico com maior número de publicações, BIG DATA & SOCIETY, teve sua

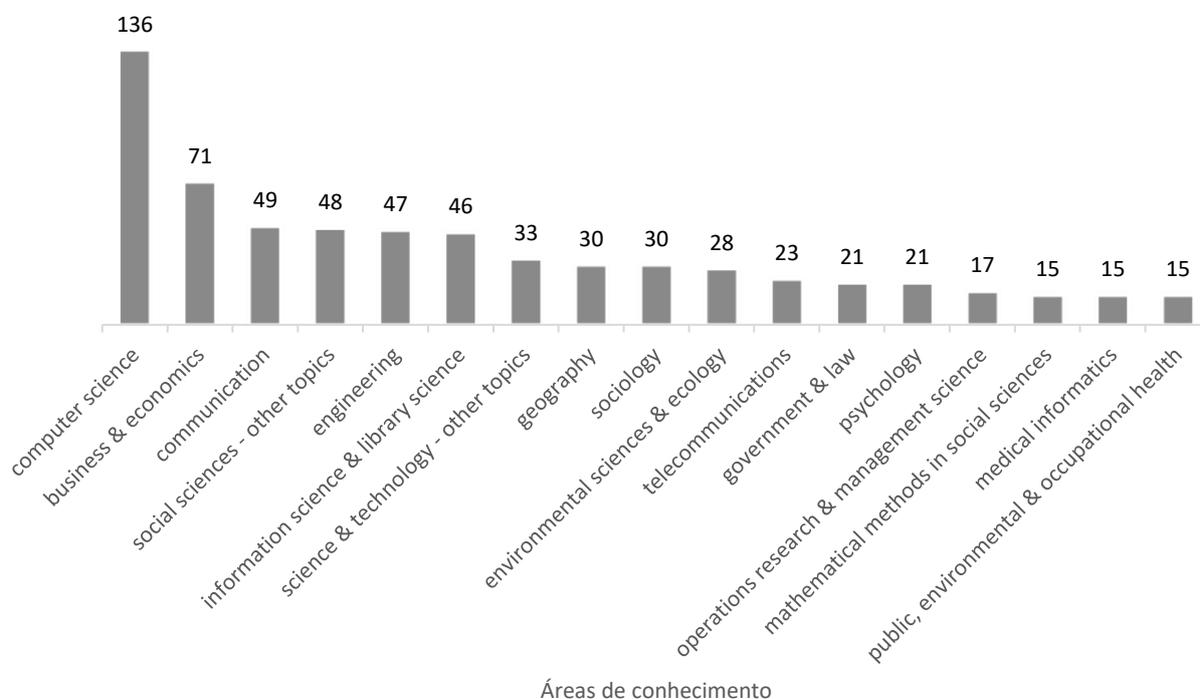
primeira edição em 2014 e se configura com um periódico de livre acesso com 2 edições por ano, por sua vez o segundo periódico com maior número de artigos publicados sobre a temática, INFORMATION COMMUNICATION & SOCIETY, possui uma edição por ano, iniciando em 1999. Estes exemplos indicam que a temática vem sendo abordada tanto em periódicos mais recentes quanto em mais tradicionais, mostrando que o interesse pela temática permeia diferentes periódicos.

**Tabela 3. Periódicos mais representativos para a temática**

Periódico	Número de artigos
BIG DATA & SOCIETY	14
INFORMATION COMMUNICATION & SOCIETY	11
PLOS ONE	10
ISPRS INTERNATIONAL JOURNAL OF GEO- INFORMATION	6
JOURNAL OF MEDICAL INTERNET RESEARCH	6
SOCIAL NETWORK ANALYSIS AND MINING	6
SUSTAINABILITY	6
BIG DATA	5
COMPUTERS IN HUMAN BEHAVIOR	5
EPJ DATA SCIENCE	5
GOVERNMENT INFORMATION QUARTERLY	5
INFORMATION FUSION	5
INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS	5

Conforme observou-se a distribuição das publicações em diferentes periódicos, buscou-se elucidar as diferentes áreas do conhecimento que estão abordando a temática. Verificou-se que os artigos estão relacionados a 72 áreas distintas. Mais da metade das áreas do conhecimento, 55%, tiveram apenas 1 ou 2 artigos publicados (31 áreas de conhecimento com apenas 1 publicação e 9 com 2 publicações). Estes números apontam, no que se refere ao seu comportamento temáticos, que há um interesse disperso em diferentes áreas do conhecimento, havendo uma concentração em 17 áreas de conhecimento que reúnem 80% dos artigos publicados. A *Ciência da Computação* concentra 17% (136 artigos), seguida pela área de *Negócios e Economia* (13% e 71 artigos) e *Comunicação* (9% e 49 artigos), *Ciências Sociais* (8,9% e 48 artigos),

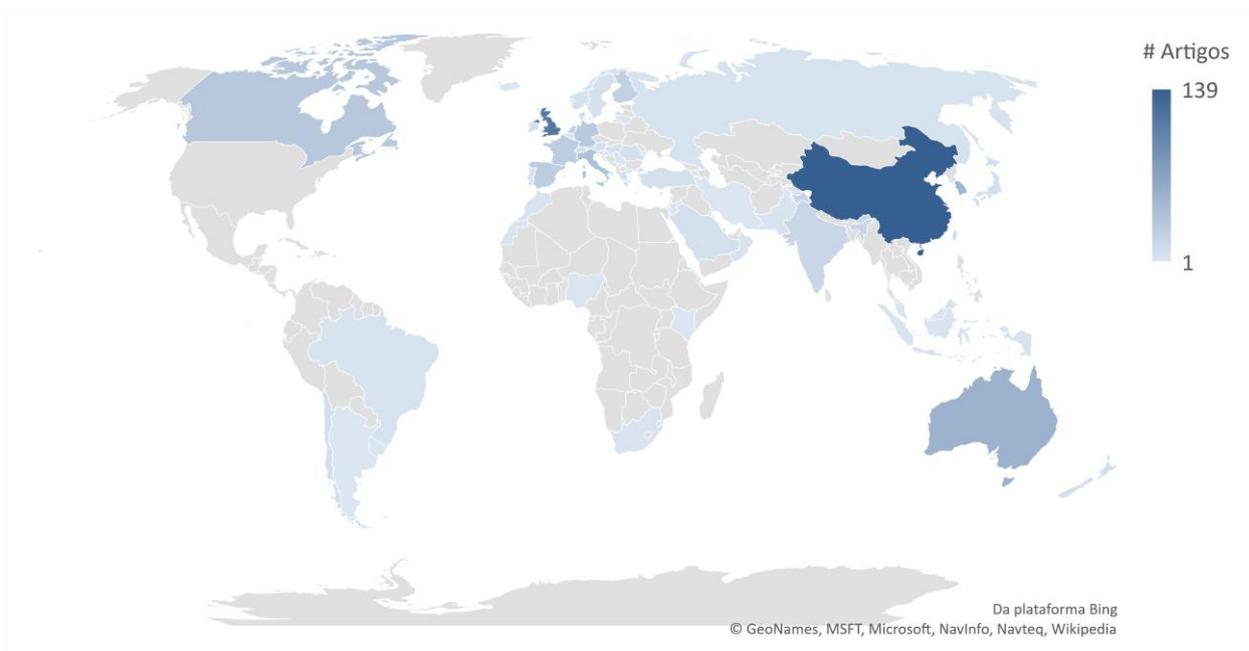
*Engenharia* (8,7% e 47 artigos) e *Ciência da Informação e Biblioteconomia* (8,5% e 46 artigos). O número de artigos publicados por cada uma das 17 áreas do conhecimento foi disposto conforme gráfico a seguir.



**Gráfico 3.** Número de publicações por área de conhecimento (17 áreas do conhecimento mais expressivas)

O potencial interdisciplinar da temática pode ser confirmado ao se verificar que as áreas do conhecimento com maior relevância possuem diferentes enfoques. Algumas possuem natureza de aspecto técnico, outras abordam mais os aspectos humanos e sociais. Este potencial foi melhor identificado ao se analisar o conteúdo dos artigos, por meio das palavras-chaves utilizadas.

Ao mesmo tempo que se observou a pulverização de interesses de diferentes áreas do conhecimento, verificou-se uma concentração geográfica quanto à nacionalidade dos autores. Dos 541 artigos publicados relacionados à temática, 72% foram realizados por pesquisadores americanos.



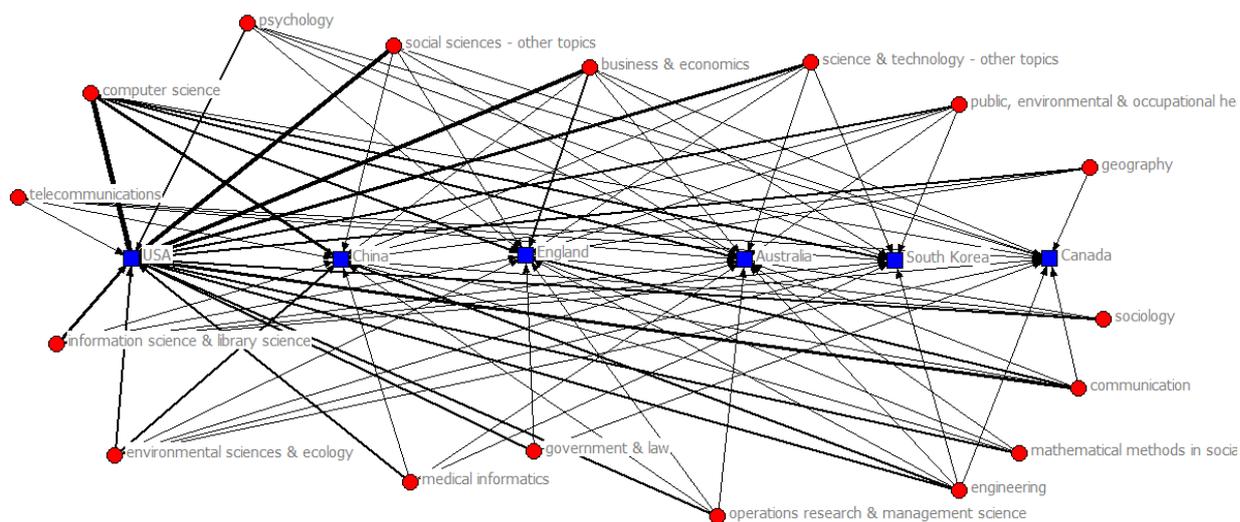
**Gráfico 4.** Distribuição geográfica dos artigos excluindo Estados Unidos

Conforme destacado no gráfico acima, no que se refere ao comportamento geográficos, a temática mostrou-se de interesse global, contando com 54 nacionalidades distintas, originárias de países dos 5 continentes. Seguido dos Estados Unidos, observou-se uma maior ocorrência da China (14%), Inglaterra (13,4%), Austrália (6,4%), Coreia do Sul (4,4%) e Canadá (4,3%) como países de origem dos autores. Por sua vez, os trabalhos foram publicados em 8 línguas distintas, sendo o inglês a língua predominante. Embora China seja o segundo país com maior número de autores nenhum artigo foi publicado considerando a língua nacional. Além do inglês, existem trabalhos publicados em espanhol (0,74%), húngaro (0,55%), alemão (0,37%), catalão (0,18%), coreano (0,18%), português (0,18%) e russo (0,18%).

Desta forma entendeu-se que há uma grande distribuição de interesses com relação às áreas do conhecimento e periódicos que realizam as publicações dos mesmos. No entanto, sob o aspecto geográfico, observou-se uma concentração de estudiosos americanos e domínio da língua inglesa.

Para melhor compreensão do comportamento da temática realizou-se uma análise mais aprofundada relacionando primeiramente estas áreas do conhecimento e os principais países de nacionalidade dos estudiosos.

Para a análise, foi considerado as 17 principais áreas de conhecimento que possuem artigos vinculados à temática e os 6 principais países de origem dos pesquisadores de forma a proporcionar um melhor entendimento sobre possíveis enfoques diferenciados. No diagrama a seguir, a espessura das linhas de conexão indica a correlação entre os itens, sendo ela mais grossa quanto maior for a correlação.



**Gráfico 5.** Diagrama de correlação de áreas do conhecimento

Dado ao maior número de trabalhos vinculados aos Estados Unidos, observou-se que este é o país com maior representatividade para todas as principais áreas do conhecimento a exceção de *Engenharia e Telecomunicações* que estão mais fortemente relacionadas à China.

Dentre as áreas do conhecimento, foi possível verificar que os artigos publicados por estudiosos dos Estados Unidos estão mais fortemente relacionados a áreas como *Ciência da Computação (11,5%)*, *Ciências Sociais (8%)*, *Negócios e Economia (8%)*, *Comunicação (6,5%)* e *Ciências e Tecnologia (6,3%)*, indicando uma abrangência de perspectiva da temática uma vez que há um destaque tanto para áreas mais técnicas quanto de caráter humano-social.

Para os demais países a *Ciência da Computação* tem um peso maior, mais de 20% dos artigos. Neste aspecto, a Inglaterra se diferencia dos demais por seus artigos estarem mais fortemente relacionados à área de *Negócios e Economia (19%)* enquanto

*Ciência da Computação* tem o mesmo peso de *Comunicação* (9,5% cada), não havendo artigos relacionado à *Telecomunicações*.

A China, por sua vez atribui maior peso às áreas mais técnicas. *Ciência da Computação* (21,5%), *Engenharia* (13%), *Ciência Ambiental e Ecologia* (7,4%), não tendo trabalhos relacionados à temática associados *Direito e Governo* e *Psicologia*.

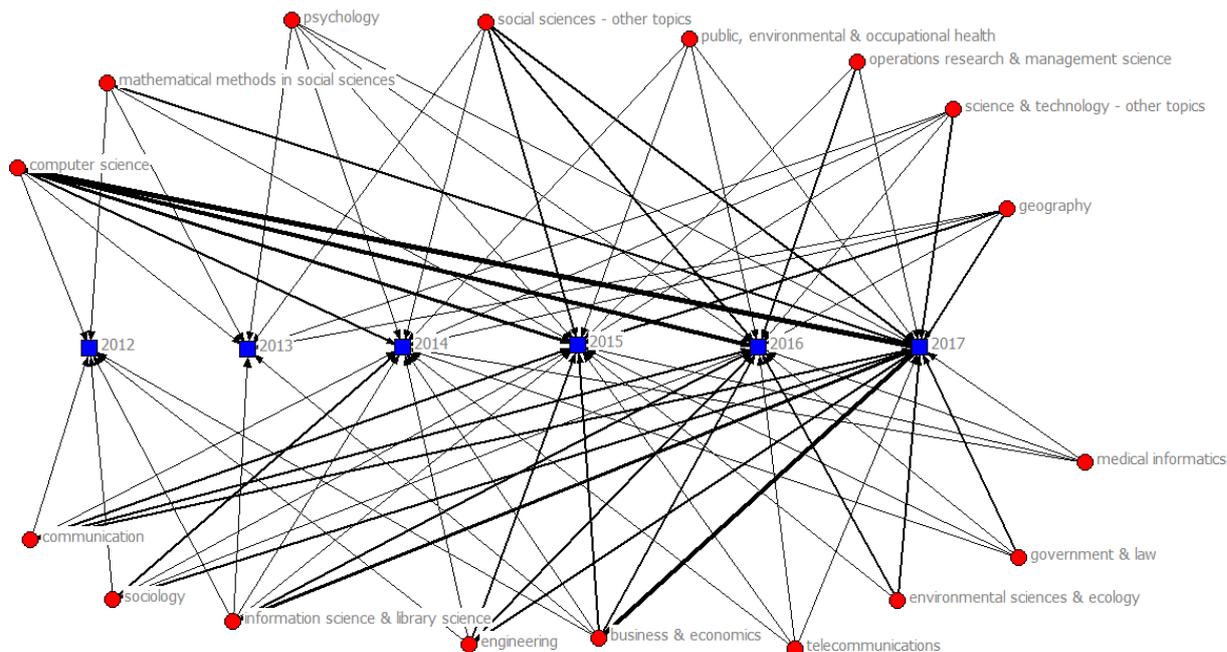
A Austrália foi o segundo país com maior número de artigos relacionados à temática vinculado à *Psicologia*, que corresponde também à terceira área de maior relevância para a temática no país. *Ciência da Computação* (23%), *Negócios e Economia* (17%) e *Psicologia* (9,6%).

A Coreia do Sul não apresenta trabalhos vinculados à *Comunicação*, *Ciência Social*, *Sociologia*, *Geografia* e *Direito e Governo*, sendo seu enfoque de estudo muito mais tecnicistas. *Ciência da Computação* (34%), *Telecomunicações* (14,6%).

As publicações sobre a temática para o Canadá estão mais vinculadas à *Ciência da Computação* (20,6%), *Ciência da Informação* e *Biblioteconomia* (15%), *Ciências Sociais* e *Engenharia* (12% cada), sendo que não há publicações relação à *Ciência e Tecnologia* e *Sociologia*.

De forma geral, quanto ao aspecto geográfico e temático, foi possível verificar que dentre os principais países há uma pluralidade de interesses das diferentes áreas do conhecimento, sendo os Estados Unidos o mais significativo para quase todas as áreas do conhecimento exceto *Engenharia* e *Telecomunicações*. Também foi possível verificar um destaque para Inglaterra no que se refere a *Negócios e Economia*, *Comunicação*, *Ciências Sociais* e *Ciências da Informação*.

Complementar ao entendimento do comportamento temático de forma geográfica, dado a sua natureza de estudos recentes, objetivou-se compreender também o seu desenvolvimento temporal, ou seja, como que os interesses das diferentes áreas do conhecimento evoluíram ao longo destes anos iniciais de estudos. Para tal, analisou-se as 17 principais áreas de conhecimento e sua distribuição ao logo dos anos iniciando em 2012



**Gráfico 6.** Diagrama de correlação de ano de publicação e área do conhecimento

Conforme o diagrama foi possível verificar que dentre as áreas do conhecimento que apresentam maior número de artigos vinculadas à tema considerando-se todo o período de publicações analisado, no início não estavam vinculadas aos primeiros artigos publicados em 2012, revelando um aumento de interesse de outras áreas sobre a temática ao longo dos anos e configurando uma temática de interesse de diversas áreas. Por sua vez também pode-se verificar que áreas pioneiras, as que estavam vinculadas aos artigos publicados em 2012, seguiram seu interesse na temática e ganharam maior relevância com o tempo, indicando assim não uma transição de interesse de áreas de conhecimento e sim uma ampliação.

Sob o aspecto da continuidade e consolidação do interesse das diversas áreas, verificou-se alguns casos de oscilação, ou seja, algumas rupturas na continuidade e aumento de publicações. Sob este aspecto destacou-se *Logística e Administração* que teve uma maior relevância no ano de 2016, porém que não se manteve no ano seguinte, e *Geografia* que teve maior relevância em 2015, perdendo destaque em 2016, mas retomando em 2017. Este comportamento sugere que para algumas áreas do conhecimento a temática ainda não está totalmente consolidada como focos de interesse de contínuo crescimento.

Para melhor compreender o comportamento temático e empírico dos artigos em estudo, verificou-se as palavras chaves destacadas pelos autores e a partir dessa análise ambicionou-se melhor compreender de que forma as áreas do conhecimento estão abordando a temática.

Foram identificadas 1735 palavras-chaves, sendo que destas, 1501 palavras foram citadas uma única vez (86,5%), ou seja, menos de 15% das palavras-chaves tiveram alguma recorrência nos trabalhos analisados.

Ao analisar as 233 (13,5%) palavras-chave que tiveram recorrência de citação (mais que 1 vez), destacaram-se: “Big Data” (260 artigos), “Social Media” (196 artigos) e “Twitter” (71 artigos), sendo que as demais palavras-chave obtiveram uma frequência até 60% menor ou mais do que a terceira mais citada. A maior presença das palavras-chave “Big Data” e “Mídias Sociais” é compreendida levando-se em consideração que a temática foi formada tendo como critério a combinação destas. As demais palavras-chaves com até 4 recorrências foram dispostas conforme ilustração a seguir, que consideram o tamanho da fonte relacionado a quantidade de aparições do termo.



Gráfico 7. Wordcloud principais palavras chaves

Conforme ilustração foi possível verificar o destaque de palavras como “Análise de Sentimentos”, “Redes Sociais”, “Mineração de Dados”, “Machine Learning” e “Facebook”. De acordo com a quantidade de citações verificou-se uma presença mais destacada para tópicos relacionados às tecnologias e ferramentas empregadas, como é o caso da Mineração de Dados (18 artigos), Machine Learning (17 artigos), Big Data Analytics (14 artigos), Mineração de Texto (11 artigos), entre outros. No entanto, questões e abordagens mais humanísticas, que fazem uma reflexão como a forma com que estes recursos estão sendo empregados e seus efeitos na sociedade também aparecem com certa relevância, como é o caso de Ética (10 artigos), Privacidade (10 artigos), Vigilância (7 artigos), Ativismo (4 artigos), Tomada de Decisões (4 artigos) e Mobilidade Urbana (4 artigos). Desta forma pode-se perceber que além dos aspectos funcionais e técnicos das ferramentas empregadas, há também um interesse pelo estudo da sua utilização.

Para melhor compreensão de como os assuntos estão sendo abordados, verificou-se a correlação das principais palavras chaves com as áreas do conhecimento dos respectivos artigos. Para a seleção das palavras chaves principais levou-se em consideração aquelas com até 5 citações, o que corresponde a 38 palavras, o equivalente a 2,5% das palavras chaves e 32,5% das recorrências.

Com o propósito de ilustrar de maneira mais fácil as correlações, utilizou-se uma tabela com o número de usos das palavras-chaves para cada área do conhecimento. Desta forma foi possível verificar que para a maioria das palavras a área do conhecimento com maior número de artigos que faz uso dela é a *Ciência da Computação* (60,5% das palavras selecionadas). Este fato pode ser compreendido pela maior quantidade de artigos vinculados a esta área do conhecimento e evidências que mesmo dentro de uma área do conhecimento há diferentes abordagens de estudo.

Embora haja uma grande quantidade de palavras chaves que foram utilizadas em artigos publicados e vinculados à *Ciência da Informação*, verificou-se que as palavras que sugerem um discussão de caráter comportamental e social do uso das mídias sociais e big data como Ética e Vigilância não foram destacadas aos artigos publicados vinculados a esta área, no entanto há ocorrência utilizando as palavras *Privacidade* e *Pesquisa Ética*, indicando que são temas também possíveis de ser explorado sob a perspectiva da área de conhecimento.

**Tabela 4. Palavras-chave por área do conhecimento**

	computer science	business & economics	communication	social sciences - other topics	engineering	information science & library science	science & technology - other topics	sociology	geography	environmental science & ecology	telecommunications	government & law	psychology	operations research & management science	mathematical methods in social sciences	medical informatics	public, environmental & occupational health
big data	65	29	25	22	24	20	13	17	18	17	10	9	9	9	7	6	10
social media	39	32	24	19	14	17	5	15	9	9	6	7	7	9	7	9	4
Twitter	14	4	8	11	8	7	3	6	7	3		6	5	6		1	2
sentiment analysis	9	5	1	1	2	2		1		1	1	1	3	1	1	1	1
social networks	4	3		2		1	2	1	2	2			2			2	
data mining	7	1	2	1	2		2			1			1				1
machine learning	4	2		1	1			1		2			1		1	1	
Facebook	1	1	7	3		3		1		1	1	2	1			1	
big data analytics	5	5	1	3		3		1							1		
computational social science	1		3	1			1	4					1		3		
social network analysis	8	1	1	1	4	2					1			2			
text mining	7	2		1	2	2					1			1		1	
ethics			2	1		2	1	3	1							1	
privacy	1		3			2			1		1						1
Hadoop	4	1			3		2				2						
Internet			2	1		1		1	1		1	3				1	
Analytics	1	2	1	1		2		1		1				1			
social media analytics	6	2		1	1	3		1						1			

crowdsourcing	3		2	3	1	1	1	2		1	1		1
data analytics	1	1		3	2		1	1		1		1	1
MapReduce	5	1							2				
business intelligence	4	2		1			1	1	1				
predictive analytics	2	2			1						1		
surveillance			2	1							1		2
Internet of Things	2	2			1								
algorithms	2	2			1	1	2						
content analysis	1	1	1		1	1	1	3		1			
Cloud computing	6				1					1			
user-generated content	1	2		1	1	2	1			1			
visual analytics	4	1		1		1							
Natural language processing	3							1					1
GIS			1				1	2	1				1
research ethics	1		1			2	1						
smart city	1	1		1	1		1			1			1
Social media analysis	2	1				1		1			1		
technology			1					1				1	
e-Government	1					2				1	1		
methodology			2		1		1						

---

Verificou-se que apesar de na perspectiva geral Twitter ser mais frequente do que o Facebook, sob o âmbito de pesquisa na área de *Comunicação* a quantidade de artigos relacionado a estas palavras é bastante próximo (8 e 7 respectivamente), sendo que para Facebook a *Comunicação* é a área de conhecimento mais relevante. Além disso verifica-se que a área despontando o interesse relacionado a vigilância e privacidade.

Observou-se as palavras chave mais predominante entre as principais áreas de conhecimento são Big Data (100%), Mídia Social (100%), Twitter (88%), Análise de sentimento (88%) e Facebook (65%).

Quando verificada a distribuição de palavras chaves associada à nacionalidade dos autores, observou-se que os Estados Unidos foi o país de maior representatividade para a maioria das principais palavras chaves, exceto para Análise de Rede Social, Algoritmo, Computação em nuvem e Smart City.

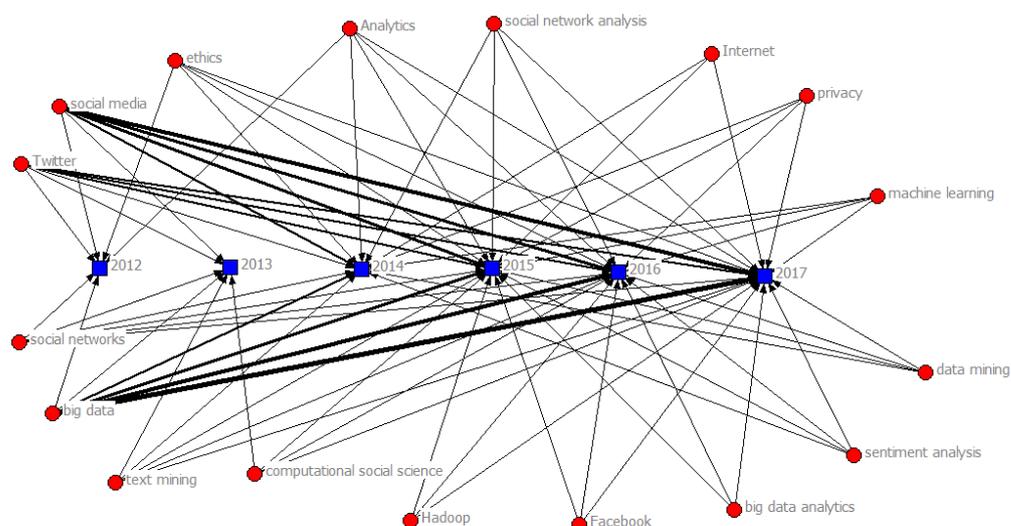
A China mostrou-se mais relevante em alguns temas como Análise de Rede Social, MapReduce, Internet das coisas e Computação na nuvem, assim como o Canadá se destaca com relação ao conteúdo gerado por usuário, smart city e análise de mídia social, enquanto as demais nacionalidades não se destacam em boa parte dos temas.

**Tabela 5. Palavra-chave por geografia**

	USA	China	England	Australia	South Korea	Canada
big data	210	38	34	15	8	6
social media	132	25	27	17	12	7
Twitter	60	3	11	6	2	6
sentiment analysis	19	2	1	3		
social networks	22	1	1	1	1	1
data mining	16	2		1	3	
machine learning	16	1		1		1
Facebook	7	1	2	2	1	1
big data analytics	15	2	2	2		
computational social science	14	2	3			
social network analysis	4	6	1			1
text mining	13	1	1	1	1	
ethics	13		1			
privacy	7	1	1	1	1	
Hadoop	2	1			1	
Internet	5		2			

Analytics	8		2			
social media analytics	7	1	1	2		
crowdsourcing	7	1				2
data analytics	11	1			1	2
MapReduce	3	3				
business intelligence	5				1	
predictive analytics	7			1		
surveillance	12				1	
Internet of Things	9	4	1			
algorithms	0	1	1	1		
content analysis	4		1			
Cloud computing	1	2			1	1
user-generated content	1	1				1
visual analytics	5	1			1	1
Natural language processing	8		1			
GIS	3	2	1			
research ethics	3				1	
smart city	1					2
Social media analysis	0	1				1
technology	8		1			1
e-Government	1	1			1	
methodology	4		1			

Para melhor compreender como estas questões evoluíram no decorrer do tempo, 65 palavras-chave mais frequentes, que obtiveram mais que 4 citações (63% das palavras) foram dispostas no diagrama a seguir de forma a permitir esta visualização.



**Gráfico 8.** Diagrama de correlação de ano de publicação e palavra-chave

Observou-se que no ano inicial das publicações, 2012, apenas 6 das 17 palavras chaves principais já estavam associadas aos artigos sobre a temática. Big Data, Mídia Social e Twitter seguiram presentes nos artigos dos anos seguintes aumentando sua relevância, enquanto que outras palavras como Rede Social, Analytics e Ética continuaram presentes, porém sem aumentar o seu destaque.

Ao logo do tempo, mais palavras foram incorporadas, tal como Privacidade, sugerindo uma preocupação mais recente para a temática. Outro fato interessante observado foi que embora Facebook seja uma rede social mais antiga, a primeira a se destacar na temática foi o Twitter. Embora a quantidade de menções seja diferente entre esta duas plataformas, quando verificado o perfil dos estudos, em ambos os casos há uma maior concentração de estudos de caso representando mais de 60% dos artigos.

Com o objetivo de compreender o comportamento temático, ademais das análises quantitativas e representações gráficas, realizou-se a análise de conteúdo de todos os artigos que compõe o corpus de estudo. Nesta análise, como define-se na literatura, foi realizada interpretações com base no descritivo textual presente nos títulos e resumos dos artigos estudados. Após a leitura os trabalhos foram agrupados em diferentes categorias de forma a melhor elucidar o seu conteúdo e permitiu uma análise mais objetiva quanto à perspectiva temática.

Verificou-se a princípio a existência de 3 macro tipos de trabalhos no que se refere ao seu enfoque de trabalho:

- *Aspectos metodológicos*: artigos que apresentam modelos para análises ou que elucidam técnicas a serem utilizadas, bem como trabalhos que comparam diferentes metodologias. Estando compreendido nesta categoria todos os trabalhos que se propuseram a elucidar o estágio de entendimento sobre a temática, servindo como um guia ou ponto inicial para futuras pesquisas. Desta forma, estudos que apresentaram o estado da arte também foram inseridos nesta categoria.

Para exemplificar esta categorização é possível destacar os seguintes títulos: “Pesquisa comparativa orientada para processos usando mídias sociais”; “Modelo de computação avançada para mídia geossocial usando análise de big data”; “Uma estrutura de CRM social baseada em multi-agentes para extração e análise”.

- *Questionamentos e reflexões*: artigos que trazem uma análise social ou consequências do uso de big data em mídias sociais. Estes trabalhos trouxeram de forma

mais evidente um questionamento sobre usos de técnicas ou aplicações, além de abrirem espaço para discussões pautados em diferentes referenciais teóricos.

Para exemplificar esta categorização é possível destacar os seguintes títulos: “Vetores no Futuro da Comunicação em Massa e Interpessoal”; “A implementação do Big Data Analytics melhora o valor de mercado das empresas?”; “A publicidade de campanha negativa estimula a comunicação incivil?”.

- *Estudos de casos*: artigos que se referem a alguma aplicação específica, restrita a uma situação em análise. São situações de aplicações de metodologias ou descobertas específicas para determinada localidade, sistema ou problema. Eles utilizam-se dos referenciais teóricos e apresentam resultados concretos do uso das metodologias.

Para exemplificar esta categorização é possível destacar os seguintes títulos: “Dados de mídia social usados na medição de serviços públicos”; “Reação global aos recentes surtos do vírus Zika”; “Monitoramento da Discriminação e Teste de HIV na Mídia Social no Brasil”.

Para os casos nos quais havia a possibilidade de dupla classificações optou-se por categorizar os artigos analisados de acordo categoria com maior representatividade no estudo apresentado. Desta forma, cada um dos 541 artigos foram classificados em apenas uma das 3 categorias, chegando-se a 310 artigos dedicado ao estudo de casos específicos, 170 artigos relacionados à apresentação e discussão de metodologias e 61 artigos abordando algum tipo de reflexão sobre a temática.

Desta forma verificou-se maior representatividade nos estudos focados em aplicações e estudos de caso, com 57,3% em várias áreas do conhecimento. Uma vez que as possibilidades de aplicações do Big Data são bastante amplas e o conteúdo informacional presente nas mídias sociais desperta o interesse de variados pesquisadores, de modo a possibilitar especificidades de aplicações na ciência.

Em complemento, verificou-se que 31,4% dos estudos voltados à apresentação ou comparação metodológica foram desenvolvidos pelos diversos campos do conhecimento na tentativa de maior elucidação metodológica sobre a temática. Entendeu-se a existência de uma carência quanto a melhor compreensão de como se utilizar os conhecimentos desenvolvidos, seja pela prematuridade da temática como pela sua aplicação e atualização.

Artigos voltados a questionamentos e reflexões sobre a temática formaram a categoria com menor representatividade de trabalhos, 11,3%, indicando uma carência de estudos voltados a este objetivo. Observou-se também que temas como Fake News, que tem ganhado destaque nas mídias principalmente devido a impacto em campanhas eleitorais ainda não estão totalmente refletidos como preocupação de estudo científico, visto que apenas 3 artigos destacaram essa abordagem em seu título ou resumo vinculados a Big Data e mídia social. Este fato pode indicar um tempo natural para que temáticas emergentes na atualidade da imprensa e demais setores da sociedade estar refletido na produção científica.

Para além do entendimento sobre o comportamento temático, este trabalho visou compreender se há uma consolidação e suas frentes de pesquisa. Para tal, verificou-se que os artigos estudados são de autoria de 1.570 pesquisadores, conferindo uma média de 2,90 autores por artigo. A maioria desses pesquisadores, 91,5%, publicaram apenas 1 único artigo sobre a temática e a quantidade máxima de artigos publicada por um mesmo autor corresponde a 6 artigos.

Desta forma, em razão da dispersão de autores publicando sobre a temática observou-se que não há consolidação de autores sobre a temática, embora alguns pesquisadores com maior destaque tenham publicado mais de 2 artigos, correspondendo a 3,7% dos autores da temática (Tabela 6).

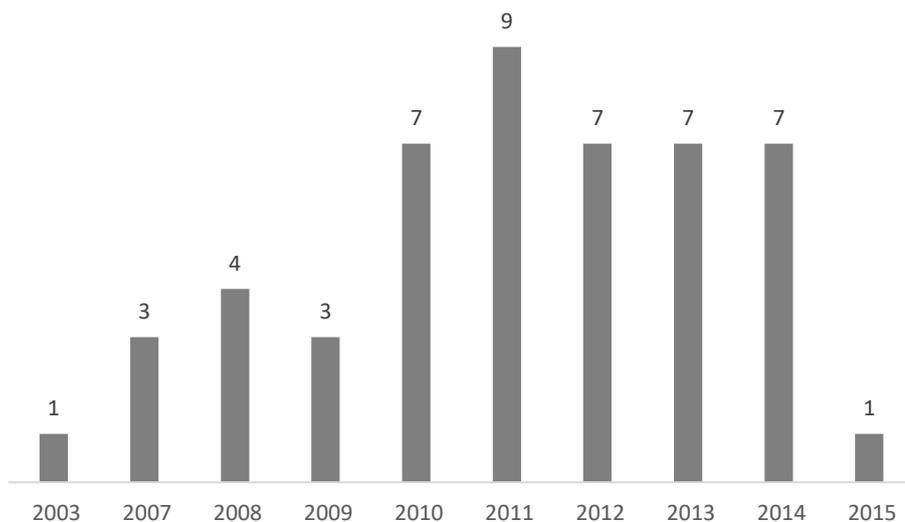
**Tabela 6. Autores mais representativos para a temática**

# Artigos	Autores
6	Poorthuis, A
5	Zook, M
4	Burnap, P
4	Cheung, M
4	Kern, ML
4	Seligman, MEP
4	She, J
4	Shelton, T
4	Williams, ML
3	Arnaboldi, M
3	Bail, CA
3	Brooker, P
3	Cao, GF
3	Crawford, K
3	Eichstaedt, JC

3	Groves, D
3	Guggenheim, L
3	Jang, SM
3	Jiang, B
3	Kosinski, M
3	Miller, WR
3	Otte, JL
3	Park, G
3	Park, HW
3	Schwartz, HA
3	Ungar, LH
3	Vargo, CJ
3	Verhoef, PC
3	Wang, W

---

Ao analisar as referências citadas por estes artigos verificou-se o total de 24.263 referências, sendo que 90,8% das referências foram citadas uma única vez. O que reforçou a diversidade de interesses sobre a temática e indicou uma falta de consolidação teórica. A referência mais utilizada está presente em 74 artigos, ou seja 13,7% dos trabalhos estudados e trata-se de um trabalho de 2012. Além dessa referência observou-se que as principais referências utilizadas datam da década atual, o que reforça atualidade da temática.

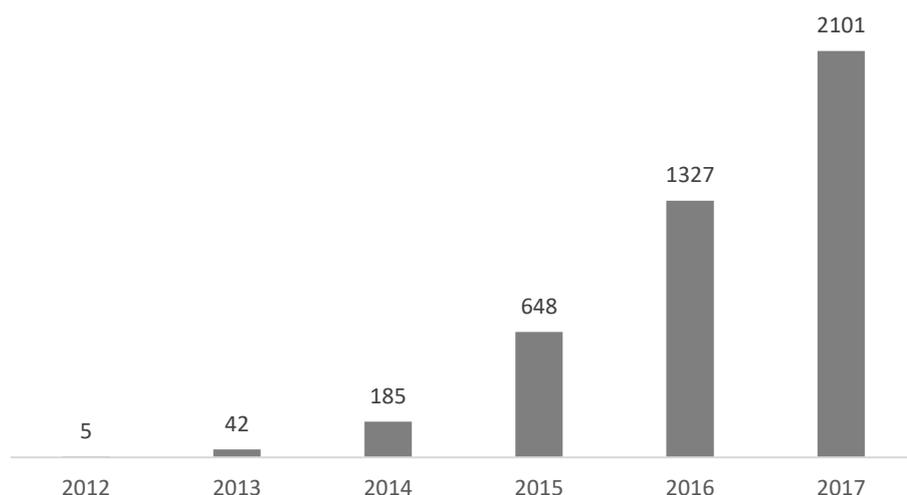


**Gráfico 9.** Ano de publicação das principais referências

Ao analisar as principais referências utilizadas pela temática, trabalhos que tiveram citação em pelo menos 10 artigos, identificou-se 50 referências que datam desde de 2003, sendo mais frequentes as publicações originárias da década atual.

Embora exista uma pluralidade de áreas do conhecimento interessadas pela temática, quando observado as principais referências utilizadas nas 17 áreas do conhecimento mais significativas para a temática, verificou-se que estas referências estão sendo compartilhadas entre as áreas do conhecimento. Ou seja, as principais referências estão servindo de base para os artigos sobre a temática, ainda que tenham enfoques de áreas de conhecimento diferentes. 56% das 50 referências mais citadas foram citadas nos artigos vinculados a mais que 50% das principais áreas do conhecimento que abordam a temática.

Conforme relatório de citações disponibilizado pela base *Web of Science*, a temática possui h-index 35, o que significa que existem 35 artigos que foram citados 35 vezes, sendo a média de citações da temática de 11,98 por item. Ao total os 541 artigos que compõem a base deste estudo foram citados 6.483 vezes, sendo que 6.166 vezes não ocorreram autocitação. Estes dados indicam a relevância da temática e como o conteúdo científico produzido vem sendo utilizado.



**Gráfico 10.** Número de citações por ano

Desta forma, a hipótese da temática estar em consolidação e se configurar como um campo emergente e em desenvolvimento pode ser confirmada ao revelar que o

material produzido vem sendo utilizado na construção do conhecimento científico sobre a temática.

Conforme mostra gráfico acima, percebeu-se que a quantidade de citações de trabalhos relacionados à temática aumentou gradativamente anualmente, indicando uma tendência de consolidação da temática.

Em se tratando das instituições ligadas às publicações, verificou-se 674 instituições, sendo que 73% destas estão ligadas a apenas um artigo, o que corroborou com a diversificação de interesses sobre a temática e a falta de frente de pesquisa definida e consolidada. Dentre às instituições com mais trabalhos associados, destacam-se 26 que possuem cada, mais de 5 artigos sobre a temática, conforme listado a seguir.

**Tabela 7. Instituições mais representativas para a temática**

# Artigos	Instituições	País
10	univ kentucky	Estados Unidos
10	univ michigan	Estados Unidos
10	wuhan univ	China
8	chinese acad sci	China
8	indiana univ	Estados Unidos
8	univ penn	Estados Unidos
7	harvard univ	Estados Unidos
7	univ manchester	Estados Unidos
7	univ melbourne	Austrália
6	northeastern univ	Estados Unidos
6	northwestern univ	Estados Unidos
6	ucl	Estados Unidos
6	univ maryland	Estados Unidos
6	univ oxford	Estados Unidos
6	univ washington	Estados Unidos
6	univ wisconsin	Estados Unidos
5	arizona state univ	Estados Unidos
5	carnegie mellon univ	Estados Unidos
5	hong kong univ sci & technol	Hong Kong
5	politecn milan	Itália
5	tsinghua univ	China
5	univ amsterdam	Holanda
5	univ calif los angeles	Estados Unidos
5	univ cambridge	Reino Unido
5	virginia tech	Estados Unidos
5	yeungnam univ	Coréia do Sul

Observou-se que 17 destas instituições com mais de 5 artigos publicados sobre a temática são universidades americanas. Isto por um lado confirmou o interesse deste país relacionado à temática conforme verificado pelo grande número de pesquisadores envolvidos, mas também revelou que internamente no país há diversas instituições abordando a temática, não havendo, portanto, um polo consolidado.

Fundamentalmente, a temática de Big Data e Mídia Social na ciência tem-se desenvolvido de forma acelerada face a quantidade de artigos publicados, a diversidade de áreas de conhecimento envolvidas, a articulação temática dos assuntos produzidos em variados canais de divulgação em curto espaço de tempo. Há multiplicidade de autores, instituições e países envolvidos no processo de consolidação da temática.

## 5. CONCLUSÃO E CONSIDERAÇÕES FINAIS

Conforme apontado por Castells (1999) e Zalis (2016) o período atual é marcado por rápidas e significativas mudanças observadas em diferentes áreas, fomentadas pelo desenvolvimento tecnológico, vinculado à disseminação da internet. Inserido neste contexto, as mídias sociais proporcionam o encontro e interações dos indivíduos no ciberespaço, permitindo o registro, a digitalização, de interações que anteriormente ocorriam majoritariamente de forma não intermediada pela internet.

Esta mudança permitiu à ciência, em especial à ciência social aplicada, uma mudança de paradigma acerca da coleta de dados, conforme evidenciado por Golder e Macy (2012). Diante da possibilidade e acesso a uma grande quantidade de dados sobre o relacionamento e interação social dos indivíduos e à sua característica de rápida atualização, entre outras características favoráveis, a pesquisa se favoreceu de forma a também ganhar mais velocidade na compreensão dos fenômenos da atualidade a qual está inserida.

Este fator se torna especialmente importante quando identificado o presente momento de contínuas mudanças evidenciadas por Bauman (2011) ao caracterizar por meio da associação com líquido as contínuas transformações e instabilidades. De forma complementar Lipovetsky (2004) destaca a intensidade empregada nas relações, sejam entre os indivíduos, seu trabalho, planos, corpo e bem-estar, entre outros, marcado pelo identificador hiper.

Sendo assim, estas percepções sobre o presente momento corroboram para elevar a importância de a ciência poder utilizar-se de recursos tecnológicos capazes de aproximá-la das questões atuais em velocidade de pesquisa que permita contribuir com a análise e compreensão desta sociedade em rápida e constante transformação.

A pesquisa partiu da hipótese da crescente utilização e aplicação de Big Data em mídias sociais confirmando-se a possibilidade de emprego sob diversas perspectivas temáticas e permitindo o desenvolvimento e aprimoramento de técnicas de utilização na obtenção de conteúdo informacional para a ciência. Pressupondo-se grande impacto de estudos desenvolvidos a partir da lógica de aplicação da ferramenta de Big Data em mídia social, além de marcos teóricos pouco desenvolvidos, ancorados em poucas áreas do conhecimento, sobretudo nas ciências exatas.

Para tal, foi apresentado panorama da produção científica que relaciona o uso de Big Data e Mídias Sociais, com base nos artigos indexados na base de dados *Web of Science*. Com caráter exploratório e de abordagem quali-quantitativa, o estudo aplicou procedimentos de análises bibliométricas e cientométricas acrescidas de análise de conteúdo. O corpus constituiu-se de 541 artigos indexados na base de dados *Web of Science* que permitiu identificar o comportamento temático e empírico à luz do posicionamento geográfico e temporal, bem como analisar as contribuições e limitações nos diferentes campos do conhecimento.

Primeiramente, verificou-se que os primeiros trabalhos sobre a temática datam de 2012, configurando-se como bastante recente, ao mesmo tempo que apresentou uma tendência de crescimento acelerada no número de artigos desde o início das publicações, indicando o processo de consolidação de interesse sobre a temática. Este comportamento sugere um reforço nas perspectivas levantadas no suporte teórico que evidencia a velocidade das transformações influenciadas pelo desenvolvimento tecnológico associado à internet.

Seu caráter recente e de interesse interdisciplinar foi constatado pela diversidade de periódicos, áreas de conhecimento, autores e referências utilizadas. Os 541 artigos analisados foram publicados em 344 periódicos distintos, o que confere uma média de 1,57 artigos por periódico, indicando uma pulverização da temática. Apenas pouco mais de 16% do total dos artigos foram divulgados por periódicos com mais de 5 publicações.

De forma similar, verificou-se a pulverização de interesses em relação às áreas do conhecimento uma vez que o corpus se relaciona a 72 áreas do conhecimento distintas cuja participação em número de artigos publicados não se mostrou muito representativa (55%, tiveram apenas 1 ou 2 artigos publicados), apontando que há um interesse disperso em diferentes áreas do conhecimento, sendo possível identificar uma sutil concentração em 17 áreas de conhecimento que reúnem 80% dos artigos publicados.

Dentre as áreas do conhecimento de maior destaque, *A Ciência da Computação* concentra 17% (136 artigos), seguida pela área de *Negócios e Economia* (13% e 71 artigos) e *Comunicação* (9% e 49 artigos), *Ciências Sociais* (8,9% e 48 artigos), *Engenharia* (8,7% e 47 artigos) e *Ciência da Informação e Biblioteconomia* (8,5% e 46 artigos). Sob este aspecto destaca-se também a presença de áreas do conhecimento ligadas a questões sociais, reforçando o interesse desta área de pesquisa.

Desta forma, o potencial interdisciplinar da temática pode ser confirmado ao se verificar que as áreas do conhecimento com maior relevância possuem diferentes enfoques quanto às abordagens e assuntos trabalhados. Algumas possuem natureza de aspecto técnico, outras abordam mais os aspectos humanos e sociais. Vale ressaltar que para um melhor entendimento sobre como as áreas do conhecimento estão enfocando suas pesquisas deve-se aprofundar nesta direção.

Ao mesmo tempo que se observou a pulverização de interesses de diferentes áreas do conhecimento, verificou-se uma concentração geográfica quanto à nacionalidade dos autores. Dos 541 artigos publicados relacionados à temática, 72% foram realizados por pesquisadores norte-americanos, sendo o inglês a língua principal para as publicações.

Embora haja essa concentração, a temática se mostrou de interesse global, contando com 54 nacionalidades distintas, originárias de países dos 5 continentes. Seguido dos Estados Unidos, observou-se uma maior ocorrência da China (14%), Inglaterra (13,4%), Austrália (6,4%), Coreia do Sul (4,4%) e Canadá (4,3%) como países de origem dos autores. Por sua vez, os trabalhos foram publicados em 8 línguas distintas, sendo o inglês a língua predominante, 97% dos trabalhos.

Conforme pressuposto de pesquisa, no qual haveria maior predomínio de estudos relacionados à área exatas, observou-se que para os demais países a *Ciência da Computação* é a área do conhecimento com maior número de artigos sobre a temática, mais de 20% dos artigos.

De forma geral foi possível verificar que dentre os principais países há uma pluralidade de interesses das diferentes áreas do conhecimento, sendo os Estados Unidos o mais significativo para quase todas as áreas do conhecimento exceto *Engenharia e Telecomunicações*. Também foi possível verificar um destaque para Inglaterra no que se refere a *Negócios e Economia, Comunicação, Ciências Sociais e Ciências da Informação*.

Observou-se que o interesse das áreas do conhecimento evoluiu no panorama temporal, indicando um aumento de interesse de outras áreas sobre a temática ao longo dos anos e configurando uma temática de interesse de diversas áreas. Por sua vez também pode-se verificar que áreas pioneiras, as que estavam vinculadas aos artigos publicados em 2012, seguiram seu interesse na temática e ganharam maior relevância

com o tempo, indicando assim não uma transição de interesse de áreas de conhecimento e sim uma ampliação e consolidação.

Com relação ao aspecto temático, foram identificadas 1735 palavras-chaves, sendo que destas, 1501 palavras foram citadas uma única vez (86,5%), ou seja, menos de 15% das palavras-chaves tiveram alguma recorrência nos trabalhos analisados, indicando também uma dispersão de assuntos relacionados e sugerindo uma possibilidade de aprofundamento de estudos para melhor compreensão acerca da motivação para diferentes temáticas.

Ao mesmo tempo, identificou-se maior representatividade nos estudos focados em aplicações e estudos de caso, com 57,3%, seguido de 31,4% dos estudos voltados à apresentação ou comparação metodológica desenvolvidos pelos diversos campos do conhecimento na tentativa de maior elucidação metodológica sobre a temática. Artigos voltados a questionamentos e reflexões sobre a temática formaram a categoria com menor representatividade de trabalhos, 11,3%, indicando uma oportunidade de estudos voltados a este objetivo. Este comportamento confirma o pressuposto de maior aplicação prática das técnicas e carência no embasamento teórico científico.

Com relação à autoria dos artigos, foram registrados 1570 pesquisadores, sendo a grande maioria desses pesquisadores, 91,5%, com apenas 1 único artigo sobre a temática e a quantidade máxima de artigos publicada por um mesmo autor correspondente a 6 artigos. Desta forma, observou-se que não há uma concentração consolidada de autores sobre a temática, embora haja alguns com maior destaque.

Ao analisar as referências citadas por estes artigos verificou-se o total de 24.263 referências, sendo que 90,8% das referências foram citadas uma única vez. O que reforça a diversidade de interesses sobre a temática e indica uma falta de consolidação teórica. A referência mais utilizada está presente em 74 artigos, ou seja 13,7% dos trabalhos estudados e trata-se de um trabalho de 2012.

Embora exista uma pluralidade de áreas do conhecimento interessadas pela temática, quando observado as principais referências utilizadas nas 17 áreas do conhecimento mais significativas para a temática, verificou-se que estas referências estão sendo compartilhadas entre as áreas do conhecimento. Ou seja, as principais referências estão servindo de base para os artigos sobre a temática, ainda que tenham enfoques de áreas de conhecimento diferentes. 56% das 50 referências mais citadas

foram citadas nos artigos vinculados a mais que 50% das principais áreas do conhecimento que abordam a temática.

Ao mesmo tempo que ainda não se configuram frentes de pesquisa, foi possível observar o processo de consolidação da temática agregando os conhecimentos trabalhados por meio da constituição das citações. Conforme relatório de citações disponibilizado pela base *Web of Science*, a temática possui h-index 35, o que significa que existem 35 artigos que foram citados 35 vezes, sendo a média de citações da temática de 11,98 por item. Ao total os 541 artigos que compõem a base deste estudo foram citados 6.483 vezes, sendo que 6.166 vezes não ocorreram autocitação. Estes dados indicam a relevância da temática e como o conteúdo científico produzido vem sendo utilizado. A quantidade de citações de trabalhos relacionados à temática aumentou gradativamente anualmente, indicando uma tendência de consolidação da temática.

Em se tratando das instituições ligadas às publicações, verificou-se 674 instituições, sendo que 73% destas estão ligadas a apenas um artigo, o que corrobora com a diversificação de interesses sobre a temática e a falta de frente de pesquisa definida e consolidada.

Desta forma, pode-se concluir pela amostragem analisada dos artigos indexados pela base de dados *Web of Science*, que a temática é bastante recente e está em consolidação. Ao mesmo tempo que demonstrou uma pluralidade de interesses, a crescente nos estudos demonstram um processo de consolidação, no qual a produção científica vem sendo utilizada para corroborar com novos estudos de forma a aprimorá-los e embasá-los. A temática se prova interdisciplinar e pertinente a análises de diferentes óticas, de forma que se identificou oportunidades de construção de expertises nos diferentes ramos de estudo.

Evidenciou-se uma preponderância de estudos americanos e uma hegemonia na língua utilizada nos artigos, o que releva aos cientistas ao mesmo tempo a necessidade de se apropriar da língua para ter acesso ao arcabouço já desenvolvido, ao mesmo tempo que dada a unicidade de linguagem permite a apropriação de todo o conjunto de estudos de forma mais acessível, o que tende a incentivar e acelerar a construção de aprimoramentos no conhecimento já estabelecido.

## 6. REFERÊNCIAS

- ALVES, T.A. *De máquinas a meios de comunicação: como interfaces auxiliam no processo de apreensão de novos aparelhos multimidiáticos*. 2017. 125f. Dissertação de Mestrado. Universidade Federal de Juiz de Fora. 2017.
- ASTROM F. *Changes in the LIS research front: time-sliced cocitation analyses of LIS journal articles, 1990-2004*. Journal American Society Information Science and Technology, v. 58, n. 7, p. 947-957, 2007.
- BARDIN, L. *Análise de Conteúdo*. Lisboa, Portugal; Edições 70, LDA, 1977.
- BARDIN, L. *Análise de Conteúdo*. Lisboa, Portugal; Edições 70, LDA, 1995.
- BARRETO, A. *O mercado de informação no Brasil*. *Revista Transinformação* V.10.n.2.p.55-67. Maio/agosto 1998.
- BARRETO, A. *A questão da informação*. São Paulo em Perspectiva. V.8.n.4 p.3-8. 1994.
- BAUMAN, Z. *Tempos precários*. São Paulo: Caderno Aliás do jornal O Estado de S. Paulo. Entrevista concedida Maria Fernanda Rodrigues. 2016.
- BAUMAN, Z. *A face humana da sociologia*. São Paulo: Caderno Sabático do jornal O Estado de S. Paulo. Entrevista concedida a Laura Greenhalgh. 2011.
- BAUMAN, Z. *Legisladores e Intérpretes*. Trad. Renato Aguiar. Rio de Janeiro: J. Zahar, 2010.
- BAUMAN, Z. *Vida para consumo*. Trad. Carlos Alberto Medeiros. Rio de Janeiro: J. Zahar, 2008.
- BAUMAN, Z. *Identidade*. Entrevista a Benedetto Vecchi. Trad. Carlos Alberto Medeiros. Rio de Janeiro: J. Zahar, 2005.
- BAUMAN, Z. *Modernidade Líquida*. Trad. Plínio Dentzien. Rio de Janeiro: J. Zahar, 2001.
- BAUMAN, Z. *Globalização. As consequências humanas*. Trad. Marcus Penchel. Rio de Janeiro: J. Zahar, 1999.
- BAUMAN, Z. *O mal-estar na pós-modernidade*. Trad. Mauro Gama. Rio de Janeiro: J. Zahar, 1998.
- BELL, Gordon. Prefácio. In: HEY, Tony; TANSLEY, Stewart; TOLLE, Kristin (Org.). *O quarto paradigma: descobertas científicas na era da e-science*. São Paulo: Oficina de Textos, 2011. p. 11- 15.
- BOYD, D.; CRAWFORD, K. *Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon*. Information, Communication & Society, Abingdon, v. 15, n. 5, p. 662-679, 2012
- CALDAS, M. S e SILVA, E. C. C., *Fundamentos e aplicação do Big Data: como tratar informações em uma sociedade de yottabytes*. Bibl. Univ., Belo Horizonte, v. 3, n. 1, p. 65-85, jan/jun. 2016.
- CASTELLS, M. *A Sociedade em rede. A era da informação: economia, sociedade e cultura*. 6ª edição, v.1, São Paulo: Paz e Terra, Manuel Castells, 1999.
- CASTELLS, M. *A galáxia da Internet: reflexões sobre a internet, os negócios e a sociedade*. Trad. Maria Luiza X. de A. Borges. Rio de Janeiro: Jorge Zahar, 2003.
- CAVANILLAS J, CURRY E, WAHLSTER W. *New Horizons for a Data-Driven Economy*. Cap1 *The Big Data Opportunity*. 2016 Disponível em: <[link.springer.com/content/pdf/10.1007%2F978-3-319-21569-3\\_1.pdf](http://link.springer.com/content/pdf/10.1007%2F978-3-319-21569-3_1.pdf)>. Acessado em 18 de março de 2018.
- CHEN, C.M.; CRIBBEN, T.; MACREDIE, R.; MORAR, S. *Visualizing and tracking the growth of competing paradigms: two case studies*. Journal American Society Information

- Science and Technology. v. 53, n. 8, p. 678-689, 2002.
- CHEN, C.M.; PAUL, R.J. *Visualizing a knowledge domain's intellectual structure*. Computer. V. 34, n. 3, p. 65-71, 2001.
- CHEN, C.M. *CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature*. Journal American Society Information Science and Technology. v. 57, n.3, 359-377, 2006.
- CHEN, C.M. *Mapping Scientific Frontiers - The Quest for Knowledge Visualization*. Springer-Verlag Berlin Heidelberg; 2003.
- CHIAVEGATTO FILHO, Alexandre Dias Porto. *Uso de big data em saúde no Brasil: perspectivas para um futuro próximo*. Epidemiol. Serv. Saúde, Brasília, v. 24, n. 2, p. 325-332, June 2015. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S2237-96222015000200325&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S2237-96222015000200325&lng=en&nrm=iso)>. Acesso em 04 Mar. 2018.
- CIALDINI, R. B. *Influence: The psychology of persuasion*. New York: Collins, 1998.
- CUKIER, Kenneth Neil; MAYER-SCHOENBERGER, Viktor Mayer-Schoernberger. *The Rise of Big Data: How It's Changing the Way We Think About the World*. Foreign Affairs, p.28-40, mai/jun. 2013. Disponível em: <http://www.foreignaffairs.com/articles/139104/kenneth-neil-cukier-and-viktor-mayer-schoenberger/the-rise-of-big-data> Acessado em: 13 de abril de 2018.
- DURKHEIM, F. *Da divisão do trabalho social*. São Paulo: Martins Fontes, 1999.
- Facebook. Disponível em: <<https://newsroom.fb.com/company-info/>> Acesso em 15 de maio de 2018
- FERRARI, Pollyana. *A força da mídia social*. São Paulo: Factash Editora, 2010.
- Forgas, J. P., & Williams, K. D. (Eds.). (2001). *Social influence: Direct and indirect processes*. The Sydney symposium of social psychology. New York, NY, US: Psychology Press.
- FORGAS, J. P., & WILLIAMS, K. D. (Eds.). (2001) *Social influence: Direct and indirect processes*. The Sydney symposium of social psychology. New York, NY, US: Psychology Press.
- GARFIELD E, Sher IH, Torpie RJ. *The use of citation data in writing the history of science*. In: Philadelphia: Institute for Scientific Information; 1964.
- GARFIELD E. *Scientography: mapping the tracks of science*. Current Contents: Soc Behav Sci. v. 7, n.45, 5-10, 1994.
- GIL, A.C. *Como elaborar projetos de pesquisa*. São Paulo: Atlas, 2002.
- GIFFINGER, Rudolf et al. *Smart cities Ranking of European medium-sized cities*. Vienna: Vienna University Of Technology, 2007.
- GOLDER, Scott A. and Michael W. Macy. 2011. "Diurnal and Seasonal Mood Vary with Work, Sleep, and Daylength across Diverse Cultures." Science 333:1878-81.
- GOLDER, Scott A. and Michael Macy. 2012. "Social Science with Social Media." ASA Footnotes 40:7.
- GLANZEL, W; SCHOEPFLIN, U. *Little scientometrics – big scientometrics... and beyond*. Scientometrics, v. 30, n. 2/3, p. 375-384, 1994.
- HEMSLEY, Jeff et al. *Social Media for Social Good or Evil: An Introduction*. Social Media and Society, Jul-Set 2018, 2018. DOI:10.1177/2056305118786719.
- HJARVARD, Stig. *A mediação da cultura e da sociedade*. São Leopoldo (RS): Unisinos, 2014, p. 13 – 127 (mediação; mediação; perspectiva teórica do autor; a

política na era da mídia).

HYODO, T. *Interação universidade-empresa: a produtividade científica dos inventores da Universidade de São Paulo*. 2010. 326f. Dissertação (Mestrado em Ciência da Informação) – Universidade de São Paulo. Escola de Comunicação e Artes, São Paulo, 2010.

Instagram. Disponível em: <<https://www.instagram.com/about/us/>> Acesso em 15 de maio de 2018

Intel IT Center. Guia de Planejamento / Big Data. 2013. Disponível em: <<https://www.intel.com.br/content/dam/www/public/lar/br/pt/documents/articles/90318386-1-por.pdf>> Acesso em 14 de junho de 2018.

JAMESON, F. *Pós-modernismo. A lógica cultural do capitalismo tardio*. Trad. Maria Elisa Cevasco. São Paulo: Ática, 2000.

JAMESON, F. *Reflexões sobre Pós-modernidade, justiça e filosofia*. In CAFÉ FILOSÓFICO. Palestra proferida no Instituto CPFL. Campinas – SP, em 26 de maio de 2011. Disponível em:

<<http://www.institutocpfl.org.br/cultura/2013/05/10/reflexoes-sobre-pos-modernidade-justica-e-filosofia-com-amartya-sen-simon-blackburn-e-fredrich-jameson/>> Acesso em 20 de mar. 2018.

KOZINETS, R. *Inno-tribes: star trek as wikimedia*. In: COVA, B; KOZINETS, R; SHANKAR, A. (Eds) *Consumer tribes*. New York: Elsevier, 2007.

KOZINETS. *Click to connect: netnography and tribal advertising*. *Journal of Advertising Research*, V. 46, p. 279-288, 2006.

KOZINETS, R. *The field behind the screen: using netnography for marketing research in online communities*. *Journal of Marketing Research*, v. 39, n. 1, p. 61-72, 2002.

KOZINETS, R. *E-Tribalized marketing?: the strategic implications of virtual communities of consumption*. *European Management Journal*, v. 17, n. 3, p. 252-264, June 1999.

LA VILLE, C.; DIONNE, J. *A construção do saber: Manual de Metodologia da Pesquisa em Ciências Humanas*. Porto Alegre: Artes Médicas. Belo Horizonte: Editora UFMG, 1999.

LEVY, P. *A inteligência coletiva*. São Paulo: Edições Loyola, 1998;

LEVY, P. *Cibercultura*. Editora 34, 1999

LICKS, R. A. *Big Data: diretrizes e técnicas para preservação da privacidade*. Dissertação de Mestrado – Universidade Católica de Brasília, 2016.

LIPOVETSKY, G.; SERROY, J. *A estetização do mundo: viver na era do capitalismo artista*. Trad. Eduardo Brandão. São Paulo: Companhia das Letras, 2015.

LIPOVETSKY, G.; SERROY, J. *Felicidade Paradoxal. Ensaio sobre a sociedade de hiperconsumo*. Trad. Maria Lúcia Machado. São Paulo: Companhia das Letras, 2007.

LIPOVETSKY, G.; SERROY, J. *A era do vazio*. São Paulo: Manole, 2005.

LIPOVETSKY, G.; SERROY, J. *Os tempos hipermodernos*. Trad. Mário Vilela. São Paulo: Barcarolla, 2004.

LYOTARD, Jean-François. *A condição pós-moderna*. Trad. José Bragança de Miranda. Lisboa: Gradiva, 1989.

LYRA, Renata Maldonado da Silva. *Consumo, Comunicação e Cidadania*. *Ciberlegenda*, n. 7, 2001.

LOPES, I.L. *Estratégia de busca na recuperação da informação: revisão da literatura*. *Ciência da Informação*, v. 31, n. 2, p. 60-71, maio/ago. 2002.

MARCON, K.; MACHADO, J. B.; CARVALHO, N. J. S. *Redes sociais e arquiteturas pedagógicas: uma experiência no Facebook*. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 23., 2012, Rio de Janeiro. Anais eletrônicos...Rio de Janeiro: CBIE, 2012. Disponível em: <[www.br-ie.org/pub/index.php/sbie/article/download/1693/1454](http://www.br-ie.org/pub/index.php/sbie/article/download/1693/1454)>. Acesso: 10 maio. 2018.

MARICATO, J.M. *Dinâmica das relações entre Ciência e Tecnologia: estudo Bibliométrico e Cientométrico de múltiplos indicadores de artigos e patentes em biodiesel*. 2010. 378f. Tese (Doutorado em Ciência da Informação) – Universidade de São Paulo. Escola de Comunicações e Artes, São Paulo, 2010.

MORRIS, S.A.; YEN, G.; WU, Z.; ASNAKE, B. *Time line visualization of research fronts*. Journal American Society Information Science and Technology, v. 54, n. 5, p. 413-422, 2003.

NESELLO, P. ; FACHINELLI, A. C. . *Big Data: O Novo Desafio Para Gestão*. Revista Inteligência Competitiva , v. 4, p. 18-38, 2014.

NEIROTTI, P., De Marco, A., Cagliano, A. C., Mangano, G., & Scorrano, F. (2014). *Current trends in Smart City initiatives: Some stylised facts*. Cities, 38, 25-36. Disponível em <http://www.sciencedirect.com/science/article/pii/S0264275113001935> Acessado em 01 de julho de 2018.

OLIVEIRA, E; et al. *Análise de conteúdo e pesquisa na área de educação*. Revista Diálogo Educacional, Curitiba, v. 4,n. 9, p.11-27, maio/ago. 2003.

PEREIRA, C.A.; BAZI, R.E.R. *Flow and social relationships of knowledge in Science, Technology and Innovation: a patentometric study os Unicamp's technological production*. Scientometrics, Amsterdam, v. 81, p. 61-72, 2009.

PEREIRA, C.A.; FUJINO, A. *A pesquisa sobre patentes na Ciência da Informação: estudo bibliométrico e cientométrico da produção científica indexada na Web of Science*. Em Questão, v. 20, p. 194-206, 2014.

PERLOFF, F. *Ken Wilber's integral theory applied to mediation*. Conflict Resolution Quarterly. Vol 28, Issue 1, 2010.

PERSSON O. *The intellectual base and research front of JASIS 1986-1990*. Journal of American Society Information Science, v. 45, n. 1, p. 31-38, 1994.

PETTIGREW KE, MCKECHNIE, L. *The Use of Theory in Information Science Research*. Journal American Society Information Science and Technology, v. 52, n. 1, p. 6273, 2001.

PRICE D.S. *Little science, big science*. New York: Columbia University Press; 1963.

PRICE D.S. *Science and technology: Distinctions and interrelationships*. In: W Gruber & G. Marquis (Eds.). Factors in the transfer of technology. Boston: MIT Press; 1969.

PRICE, D.S. *Networks of scientific papers: the pattern of bibliographic references indicates the nature of the scientific research front*. Science, v. 149, n. 3683, p. 510- 515. 1965. Qzone. Disponível em: <<https://expandedramblings.com/index.php/qzone-stats-facts/>> Acesso em 15 de maio de 2018.

RHEINGOLD, H. *The Virtual Community: Homesteading on the Electronic Frontier*. Reading, Massachusetts: Addison-Wesley. 1993.

RIDINGS, C. M., GEFEN, D., & ARINZE, B. *Some antecedents and effects of trust in virtual communities*. Journal of Strategic Information Systems, 11 (3–4), 271–295. (2002).

RIDINGS, C. M., GEFEN, D., *Virtual community attraction: Why people hang out online*. Journal of Computer-Mediated Communication, 10(1). (2004).

RÖDDER, S; SCHÄFER, MS. *Repercussion and resistance. An empirical study on the interrelation between science and mass media*. 2010. Communications 35 (3), 249-267

- RÜDIGER, F. (Org.). *As teorias da cibercultura: perspectivas, questões e autores*. Porto Alegre: 2ª edição, Sulina, 2013.
- SANTAELLA, Lucia. *Culturas e Artes do Pós-Humano*. Da cultura das mídias à cibercultura. 1. ed. São Paulo: Paulus, 2003. v. 1. 334p
- SMALL, H. *Co-citation context analysis and the structure of paradigms*. Journal of Documentation. v. 36, p. 183-196, 1980.
- SMALL, H. *Co-citation model of a scientific speciality: a longitudinal study of collagen research*. Social Studies of Science. v. 7, p. 139-166, 1977.
- SOUSA, S. R. e FRANCISCO, A. L., *O Método da Cartografia em Pesquisa Qualitativa: Estabelecendo Princípios...Desenhando Caminhos. Investigação Qualitativa em Saúde*. Vol.2. 2016
- Tenkanen H, Minin E, Heikinheimo V, Hausmann A, Herbst M et. al. *Instagram, Flickr, or Twitter: Assessing the usability of social media data for visitor monitoring in protected áreas*. Scientific Reports | 7: 17615 | DOI:10.1038/s41598-017-18007-4, 14 de Dezembro de 2017. Disponível em: <<https://www-nature.ez128.periodicos.capes.gov.br/articles/s41598-017-18007-4.pdf>> Acesso em 10 de abril de 2018.
- TAURION, Cezar. *Big data*. Rio de Janeiro: Brasport Livros e Multimídia Ltda., 2013.
- TESTA, J. *A base de dados ISI e seu processo de seleção de revistas*. Ciência da Informação, v. 27, n. 2., p. 233-235, maio/ago., 1998.
- TOFFLER, Alvin. *The third wave*. New York: William Morrow, 1980. 544 p.
- TURNER, V., Gantz, J. F., Reinsel, D., & Minton, S. (2014). *The digital universe of opportunities: rich data and the increasing value of the internet of things*. Rep. from IDC EMC.
- Twitter. Disponível em: <<https://www.omnicoreagency.com/twitter-statistics/>> Acesso em 15 de maio de 2018
- VARGAS-QUESADA B.; MOYA-ANEGÓN, F.; CHINCHILLA-RODRÍGUEZ, Z., GONZÁLEZ-MOLINA, A. *Showing the essential science structure of a scientific domain and its evolution*. Information Visualization, v. 9, n. 4, p. 288-300, 2010.
- VILAJOSANA, Ignasi & LLOSA, Jordi & MARTINEZ, Borja & DOMINGO PRIETO, Marc & ANGLES, Albert & VILAJOSANA, Xavier. (2013). *Bootstrapping Smart Cities through a Self-Sustainable Model Based on Big Data Flows*. Communications Magazine, IEEE. 51. 128-134.
- ZALIS, P. *Um Turbilhão vem aí*. Veja, São Paulo, SP, ed. 2502, ano 49, n. 44, p.15-19, nov. 2016.
- WHITE, H.D.; MCCAIN, K. *Visualizing a discipline: An author co-citation analysis of Information Science, 1972-1995*. Journal American Society Information Science, v. 49, n. 4, p. 327-355, 1998.
- You Tube. Disponível em: <<https://www.youtube.com/intl/pt-BR/yt/about/press/>> Acesso em 15 de maio de 2018
- Weibo. Disponível em: <<https://expandedramblings.com/index.php/weibo-user-statistics/>> Acesso em 15 de maio de 2018