

UNIVERSIDADE DO VALE DO RIO DOS SINOS – UNISINOS
UNIDADE ACADÊMICA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM GESTÃO E NEGÓCIOS
NÍVEL MESTRADO

MARCELO VIGHI FERNANDES

**O FENÔMENO *BLOCKCHAIN* NA PERSPECTIVA DA ESTRATÉGIA
TECNOLÓGICA: UMA ANÁLISE DE CONTEÚDO POR MEIO DA DESCOBERTA
DE CONHECIMENTO EM TEXTO**

Porto Alegre

2018

MARCELO VIGHI FERNANDES

**O FENÔMENO *BLOCKCHAIN* NA PERSPECTIVA DA ESTRATÉGIA
TECNOLÓGICA: UMA ANÁLISE DE CONTEÚDO POR MEIO DA DESCOBERTA
DE CONHECIMENTO EM TEXTO**

Dissertação apresentada como requisito parcial para obtenção do título de Mestre em Gestão e Negócios pelo Programa de Pós-Graduação em Gestão e Negócios da Universidade do Vale do Rio dos Sinos (Unisinos).

Orientador: Prof. Dr. Jorge Verschoore

Porto Alegre

2018

F363f

Fernandes, Marcelo Vighi

O fenômeno blockchain na perspectiva da estratégia tecnológica : uma análise de conteúdo por meio da descoberta de conhecimento em texto / por Marcelo Vighi Fernandes. – 2018.
82 f. : il. ; 30 cm.

Dissertação (mestrado) — Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Gestão e Negócios, Porto Alegre, RS, 2018.
“Orientador: Dr. Jorge Verschoore”.

1. Estratégia tecnológica. 2. Blockchain. 3. Descoberta de conhecimento em texto. 4. Mineração de textos. I. Título.

CDU: 658:004

AGRADECIMENTOS

Aos meus colegas da turma do MPGN/2016-2, pela grande união que temos. Sem vocês, esta conquista não seria possível.

À Unisinos, na figura da coordenação do curso e seus professores, por oportunizar um curso de excelente nível, no qual tive experiências únicas e muito enriquecedoras.

Aos meus colegas de trabalho e ao SICREDI, por todo apoio que tive para poder realizar este mestrado.

Ao meu orientador Jorge Verschoore, por sua paciência e por todos os conselhos e orientações certeiras que muitas vezes me colocaram de volta ao caminho correto.

Aos meus pais, que sempre me ensinaram a importância da educação e que sempre investiram no meu desenvolvimento pessoal e profissional.

À minha esposa Ana Rosa e ao meu filho Angelo, que mesmo com toda a minha ausência durante os estudos, seguraram as pontas e me incentivaram sempre nesta minha busca pelo conhecimento.

Muito obrigado!!

RESUMO

A revolução das Tecnologias de Informação e Comunicação (TIC) fez as empresas perceberem a importância da estratégia tecnológica para a sua sobrevivência. *Blockchain* é uma tecnologia descentralizada de gerenciamento de transações e dados desenvolvida, primeiramente, para a moeda digital *bitcoin*. O interesse na tecnologia *blockchain* tem aumentado desde que o termo foi cunhado. Esse interesse fez com que este fenômeno se tornasse, atualmente, um dos principais tópicos de pesquisa e publicação na *Web*. O objetivo principal deste trabalho é entender de que forma o fenômeno *blockchain* está impactando na estratégia tecnológica. Para tanto, foi realizado um estudo exploratório utilizando o processo de Descoberta de Conhecimento em Texto (DCT), com a utilização de ferramentas de mineração de textos, de forma a coletar e analisar o conteúdo de um conjunto de notícias publicadas na *Web* sobre a tecnologia *blockchain*. Foram extraídas 2.605 notícias da *Web* sobre *blockchain*, publicadas entre os anos 2015 e 2017, no idioma inglês. Como resultado do estudo, foram geradas 6 proposições, mostrando que este fenômeno está impactando a estratégia tecnológica da indústria financeira direcionando o foco deste setor para implementação de soluções em arquiteturas descentralizadas. Também foi verificado que o foco estratégico tecnológico das empresas impulsionou o desenvolvimento das tecnologias de *blockchain* privadas. Identificou-se, também, os benefícios trazidos por esta tecnologia para sistemas de pagamentos entre países, diminuindo os intermediários e melhorando os processos. Ainda, foi possível mapear que esta tecnologia tem potencial para afetar as transações através de uma plataforma eletrônica comum. Em relação ao grau de maturidade desta tecnologia, foi realizada uma discussão dos achados das análises das notícias com a teoria da difusão da inovação e concluiu-se que esta tecnologia está no limiar entre as categorias de *Innovators* e *Early Adopters*. O mapa produzido por esta pesquisa ajudará empresas e profissionais na identificação de oportunidades de direcionamento das suas estratégias tecnológicas para a tecnologia de *blockchain*.

Palavras-chave: Estratégia tecnológica. *Blockchain*. Descoberta de conhecimento em texto. Mineração de textos.

ABSTRACT

The Information and Communication Technologies (ICT) revolution made companies realize the importance of technology strategy for their survival. Blockchain is a decentralized transaction and data management technology first developed for the bitcoin digital currency. The interest in blockchain technology has increased since the idea was coined. This interest has made this phenomenon one of the main topics of research and publication on the Web. The main objective of this paper is to understand how the blockchain phenomenon is impacting technology strategy. To do so, an exploratory study was conducted using the Knowledge Discovery in Text (KDT) process, with the use of text mining tools, to collect and analyze the contents of a set of news published on the Web about blockchain technology. At total, 2605 blockchain web news were extracted, all news were published between the years of 2015 and 2017, in the English language. As a result of the study, 6 propositions were generated, in which the results showed that this phenomenon is impacting the technology strategy of the financial industry, directing the focus of this sector to the implementation of solutions using decentralized architectures. It was also verified that the companies' strategic technological focus boosted the development of private blockchain technologies. Additionally, was identified the benefits brought by this technology to cross-border payment systems, reducing intermediaries and improving processes. Also, it was possible to map out that this technology has the potential to affect the transactions through a common electronic platform. In relation to the degree of maturity of this technology, a discussion of the findings with the theory of the diffusion of innovation was made and it is concluded that this technology is in the threshold between the categories of Innovators and Early Adopters. The map produced by this research will help companies and professionals in identifying opportunities to target their technology strategies to blockchain technology.

Keywords: Technology strategy. Blockchain. Knowledge discovery in text. Text mining.

LISTA DE ILUSTRAÇÕES

Figura 1 – Tela principal da pesquisa no <i>Google News</i> no idioma Inglês	30
Figura 2 – Configuração de range de data da coleta e remoção de notícias duplicadas, referente ao ano de 2015	31
Figura 3 – Tela de configuração para mostrar 100 resultados por página de pesquisa	31
Figura 4 – Evolução da pesquisa pelo termo <i>blockchain</i> entre 2015 e 2017 (<i>Google Trends</i>)	32
Gráfico 1 – Palavras mais frequentes	39
Figura 5 – Nuvem de palavras geral	40
Gráfico 2 – Palavras mais frequentes por ano	42
Figura 6 – Nuvem de comparação por ano	43
Gráfico 3 – Mapa perceptual dos termos por ano	44
Gráfico 4 – Análise de associação entre os termos mais frequentes	49
Figura 7 – Dendograma de agrupamento das palavras	52
Figura 8 – Dendograma de agrupamento dos textos	53
Figura 9 – Nuvem de comparação entre os grupos	54
Gráfico 5 – Mapa perceptual dos termos por grupo	55
Quadro 1 – Proposições.....	67

LISTA DE TABELAS

Tabela 1 – Palavras mais frequentes	38
Tabela 2 – Palavras mais frequentes por ano	41
Tabela 3 – Palavras mais frequentes dos Sentimentos Positivos e Negativos	46
Tabela 4 – Análise de associação entre os termos mais frequentes.....	47
Tabela 5 – Análise de associação entre os termos específicos	50
Tabela 6 – Análise de associação entre os Grupos e os anos de publicação.....	56

LISTA DE SIGLAS

AC	Análise de Correspondência
DCBD	Descoberta de Conhecimento em Base de Dados
DCT	Descoberta de Conhecimento em Textos
IoT	<i>Internet of Things</i>
P&D	Pesquisa e Desenvolvimento
PoC	<i>Proof of Concept</i>
PLN	Processamento de Linguagem Natural
TIC	Tecnologias de Informação e Comunicação
URL	<i>Uniform Resource Locator</i>

SUMÁRIO

1 INTRODUÇÃO	9
1.1 DEFINIÇÃO DO PROBLEMA DE PESQUISA	13
1.2 OBJETIVOS	15
1.2.1 Objetivo Geral	15
1.2.2 Objetivos Específicos	15
1.3 JUSTIFICATIVA	15
2 FUNDAMENTAÇÃO TEÓRICA	17
2.1 <i>BLOCKCHAIN</i>	17
2.2 ESTRATÉGIA TECNOLÓGICA	21
3 MÉTODO	26
3.1 DELINEAMENTO DA PESQUISA	26
3.1.1 Descoberta de Conhecimento em Textos	26
3.2 PROCEDIMENTOS DE COLETA DE DADOS	30
3.3 PROCEDIMENTOS DE ANÁLISE DE DADOS	33
4 ANÁLISE DOS RESULTADOS	37
4.1 APRESENTAÇÃO DOS RESULTADOS	37
4.1.1 Análise geral	37
4.1.2 Análise por ano	40
4.1.3 Análise de sentimentos	44
4.1.4 Análise de associação	46
4.1.5 Análise de agrupamento	51
4.1.5.1 Análise por palavra	51
4.1.5.2 Análise por arquivos de texto	53
5 DISCUSSÃO DOS RESULTADOS	57
6 CONSIDERAÇÕES FINAIS	68
6.1 LIMITAÇÕES DA PESQUISA	70
6.2 SUGESTÕES PARA TRABALHOS FUTUROS	71
REFERÊNCIAS	72
ANEXO A – CÓDIGO <i>JAVASCRIPT</i> DO GOOGLE SERPS <i>EXTRACTOR</i>	78
APÊNDICE A – PLANILHA DE URLS DO ANO DE 2015	79
APÊNDICE B – PLANILHA DE URLS DO ANO DE 2016	80
APÊNDICE C – PLANILHA DE URLS DO ANO DE 2017	81

1 INTRODUÇÃO

A revolução das Tecnologias de Informação e Comunicação (TIC) possibilitou uma grande evolução nas funcionalidades de comunicação e de compartilhamento de informações. Essa evolução fez as empresas perceberem a importância da estratégia tecnológica para a sua sobrevivência.

É geralmente aceito que existem duas abordagens para a estratégia tecnológica. Na abordagem tradicional, as companhias procuram por tecnologias que ajudam a manter ou atacar um mercado específico. Na abordagem baseada em competência, a análise é invertida de acordo com as suas competências existentes, para então a empresa buscar novas aplicações e mercados. A estratégia tecnológica deve ser coerente com a estratégia global da empresa, ou seja, as escolhas resultantes das estratégias tecnológicas e de negócios precisam ser compatíveis e se reforçarem mutuamente. Além disso, a estratégia tecnológica considera aspectos como priorização de mercados, estratégias de produtos a serem seguidas, fusões e aquisições de interesse, entre outros.

A estratégia tecnológica tem a capacidade de integrar as unidades de negócios estratégicos das empresas de forma coerente, contribuindo para o sucesso da implementação da estratégia corporativa. Por outro lado, as estratégias corporativas são influenciadas pelas inovações tecnológicas, uma vez que elas determinam o surgimento de novos produtos, serviços e processos.

Conforme Porter (1985), esta efetiva exploração dos recursos tecnológicos ajuda a empresa a construir uma vantagem competitiva sustentável capaz de garantir a sua performance ao longo do tempo. Com isso, a evolução dos recursos tecnológicos, internos e externos à empresa, tornaram-se itens da agenda estratégica das organizações.

As primeiras quatro décadas da Internet trouxe o *e-mail*, a *World Wide Web*, as *dot-coms*, as redes sociais, a *Web* móvel, o *Big Data*, a computação em nuvem e o surgimento da Internet das Coisas (*Internet of Things*). Isso reduziu os custos de pesquisa, colaboração e troca de informações. Reduziu, também, as barreiras para a entrada para novas mídias e entretenimento, surgiram novas formas de varejo e organização de trabalho, além de empreendimentos digitais sem precedentes. Através da tecnologia de sensores, a inteligência foi colocada nas carteiras, nas roupas, nos automóveis, nos edifícios, nas cidades e, até mesmo, na biologia. Ainda assim,

mesmo com o advento da Internet, sua estrutura de custos exclui cerca de 2,5 bilhões de pessoas do sistema financeiro global.

No auge da crise financeira em 2008, alguém usando o pseudônimo Satoshi Nakamoto (2008) lançou uma moeda digital chamada *bitcoin*. O *bitcoin* representou uma ideia nova no mundo financeiro pois, ao contrário de moedas como dólar ou iene, o *bitcoin* é uma moeda digital não afiliada a uma nação. Qualquer pessoa pode abrir uma carteira e receber *bitcoins* sem fornecer qualquer informação de identificação. As transações são registradas em um livro contábil universal, aberto e distribuído, visível para todos. Esse livro de registro distribuído é chamado de *blockchain*.

No sistema de *blockchain*, o livro de registro é replicado em uma imensa quantidade de bancos de dados idênticos. Cada um desses bancos fica hospedado e mantido pelos participantes da rede. Quando as alterações são inseridas em uma cópia do banco de dados, todas as outras cópias são atualizadas simultaneamente. Com isso, conforme as transações ocorrem, os registros do valor e dos ativos trocados são permanentemente inseridos em todos os livros de registro.

O processo para garantir a confiabilidade através de intermediadores é substituído quando se entende melhor a tecnologia *blockchain* pois, com o seu uso, pode-se consultar o status da transação em tempo real. Em outras palavras, em vez de confiar que uma transação será realizada conforme acordado, agora pode-se ver o status da transação e saber o que está acontecendo. No entanto, a *blockchain* não suporta apenas transações financeiras, mas também pode suportar diversos tipos de acordos. Assim, muitos dos tipos de sistemas econômicos, ou mais especificamente, comércio de direitos de propriedade, beneficiam-se de um sistema de transação baseado em confiança, seguro e transparente. No caso do *bitcoin*, o processo ocorre de forma que uma transação só poderá ser adicionada à *blockchain* quando ele inclui uma solução para um problema matemático específico. Esses problemas matemáticos são projetados para serem computacionalmente difíceis e demorados para resolver, mas simples de verificar. Isso significa que a criação de transações de criptomoeda ou, em outras palavras, a confirmação de transações antes de serem adicionadas à *blockchain*, é por padrão difícil e dispendiosa.

Saindo do contexto do *bitcoin*, existem muitos problemas que uma *blockchain* pode ajudar a resolver pois, na sua essência, uma *blockchain* armazena transações ou outros dados através do tempo. Os registros em uma *blockchain* podem ser efetivamente indelévels usando criptografia. Além disso, uma *blockchain* pode ser

projetada para ser operada por um grupo de empresas ou indivíduos em conjunto, de modo que nenhuma entidade única controle o sistema ou seus dados.

De acordo com o MIT Technology Review (SIMONITE, 2016), as tecnologias de *blockchain* vêm possibilitando que as empresas trabalhem juntas mais facilmente. O argumento é que elas podem juntar seus dados com segurança em um sistema neutro compartilhado, em vez de mantê-los trancados dentro de sistemas internos privados. Um levantamento sobre as instituições financeiras feito pela Greenwich Associates (2016) apontou que o setor investiu mais de um bilhão de dólares em iniciativas de *blockchain* no ano de 2016. As *startups* e os bancos também estão explorando um conceito conhecido como *Smart Contracts*¹, no qual as atualizações de uma *blockchain* podem ser adicionadas através de programas de computador simples – por exemplo, para fazer automaticamente um pagamento quando uma determinada transação ocorre.

Os *Smart Contracts* podem realizar operações como execução de transações financeiras ou autenticação de documentos em um acordo legal. Pretendendo substituir os administradores terceiros, como os advogados, que são exigidos em muitas transações e acordos, eles podem verificar informações e manter ou usar fundos usando criptografia semelhante àquela que sustenta o *bitcoin*. Durante o processo de compra de um bem, existe a necessidade de que um notário formule um contrato entre as partes interessadas, o que significa que há um terceiro envolvido na transferência legal, o que leva muito tempo e custa muito dinheiro. A causa disso é que os contratos precisam ser assinados por pessoas autorizadas e verificados por terceiros autorizados e confiáveis. Esse processo deve ser feito de acordo com a lei. O processo completo leva muito tempo e dinheiro e pode ser feito de forma mais rápida, se feito de outra maneira. A tecnologia baseada em *blockchain* chamada *ethereum*² oferece uma alternativa para resolver essas questões, através do uso dos *Smart Contracts*.

Segundo Beck et al. (2016), para que seja possível implementar esta ampla gama de aplicações, é necessário ter uma *blockchain* que seja projetada para suportar

¹ Um *software* (descentralizado) capaz de gerenciar contratos legais de forma autônoma, uma ideia proposta pela primeira vez por Nick Szabo, em 1997 (DUPONT; MAURER, 2015).

² *ethereum* é uma plataforma descentralizada que executa contratos inteligentes. São aplicativos que funcionam exatamente como programado sem qualquer possibilidade de tempo de inatividade, censura, fraude ou interferência de terceiros (ETHEREUM, 2018).

transações em um nível genérico através do uso de *Smart Contracts*. O *ethereum* possui uma *blockchain* completamente programável que distribui uma lógica que, normalmente, seria executada em um servidor centralizado. Um exemplo de uso dessa tecnologia se dá por equipamentos inteligentes que são chamados de *Internet of Things* (IoT). O próximo passo no desenvolvimento seria a comunicação entre esses objetos inteligentes, uns com os outros, através da Internet para, essencialmente, manter-se em funcionamento. Para organizar as tecnologias de *IoT*, empresas como a IBM e a Samsung construíram sua própria versão de *IoT* com base na *blockchain ethereum* para que os dispositivos possam interagir entre si através de *Smart Contracts* (BECK et al., 2016). Segundo Panikkar (2015), espera-se que essas transações entre dispositivos inteligentes desenvolvam o conceito que está sendo chamado de *Economy of Things*, onde todos esses dispositivos inteligentes serão um ponto de transação e de criação de valor econômico para proprietários e usuários.

De acordo com Plansky, O'Donnell e Richards (2016), empresas como HP, Microsoft, IBM e Intel já manifestam interesse em *blockchain*. No setor de serviços financeiros, algumas grandes empresas estão criando parcerias com *startups* focadas em tecnologia para explorar suas próprias possibilidades. Por exemplo, a R3, uma empresa de tecnologia financeira, anunciou em outubro de 2015 que 25 bancos haviam ingressado em seu consórcio para o desenvolvendo uma plataforma comum baseada em *blockchain*. Dentre os participantes, estavam bancos influentes como o Citi, o *Bank of America*, o HSBC, o *Deutsche Bank*, o *Morgan Stanley*, o UniCredit, o *Société Générale*, o *Mitsubishi UFG Financial Group*, o *National Australia Bank* e o *Royal Bank of Canada*. Outro experimento interessante foi feito pela Nasdaq, cujo CEO, Robert Greifeld, introduziu o Nasdaq Linq, um livro digital baseado em *blockchain* para a transferência de ações de empresas de capital fechado, também em outubro de 2015.

Se experimentos como esses forem implementados, a tecnologia *blockchain* pode se tornar uma força revolucionária em qualquer lugar onde uma negociação ocorra, onde a confiança é preciosa e onde as pessoas precisam de proteção contra roubo de identidade, incluindo o setor público, de saúde (manter registros anônimos, mas facilmente disponíveis), de varejo (lidar com compras de grande valor, como locação de veículos e imóveis) e, é claro, todas as formas de serviços financeiros.

No entanto, mesmo que o fenômeno *blockchain* pareça ser uma solução adequada para realizar transações em diversas situações, o mesmo ainda possui

desafios e limitações técnicas que precisam ser estudadas e abordadas (YLI-HUUMO et al., 2016). Ainda há lacunas no conhecimento deste fenômeno, pois existem poucos estudos que listem os principais usos destas tecnologias, os seus benefícios e quais são os problemas que ela poderá resolver. Com isso, às corporações estão em uma fase muito focada nos testes destas tecnologias associadas ao fenômeno em diversas situações e casos de uso, mas ainda sem a clareza de que isso poderá realmente resolver problemas.

Em consequência da imaturidade dessas tecnologias, faz-se importante a realização de uma pesquisa aprofundada sobre o fenômeno pois, com isso, pode-se entender melhor o real impacto do fenômeno em relação a estratégia tecnológica. Ainda, é importante salientar que não existe um padrão ou protocolo único para estas tecnologias, o que dificulta ainda mais a sua evolução quanto aos casos de uso.

1.1 DEFINIÇÃO DO PROBLEMA DE PESQUISA

A tecnologia fez surgir novas possibilidades para às organizações, assim como possibilitou, em diversas áreas, a criação de novos produtos e serviços, principalmente quando a tecnologia passou a ter um status estratégico dentro destas organizações, notadamente nos anos 1980. Nesta década, surgiram diversas indústrias baseadas em tecnologia, como telecomunicações, indústria farmacêutica, indústria química, indústria aeroespacial e biotecnologia. Em razão da grande concorrência que surgiu, as organizações perceberam a necessidade de buscar novas fontes de vantagem competitiva que permitissem que tais organizações se mantivessem no mercado. Com isso, encontraram na tecnologia o elemento que permitiu reestruturar os seus modelos de negócio.

Manter os registros contínuos das transações é uma função básica para qualquer empresa. Esses registros servem para rastrear as ações passadas e, também, guiam o planejamento das empresas para o futuro. Eles refletem o modo como a organização trabalha internamente, mas também reflete as suas relações externas. Toda empresa mantém seus próprios registros, que são privados. A grande maioria das empresas não possui um banco de dados principal de todas as suas atividades e registros. Pelo contrário, os registros são distribuídos entre sistemas diferentes, tornando o acesso aos mesmos lento e complexo.

Uma transação típica de ações, por exemplo, pode ser executada em uma fração de segundo, muitas vezes sem intervenção humana. No entanto, a transferência de propriedade das ações pode levar até uma semana para ser finalizada. Isso ocorre porque os envolvidos não possuem acesso mútuo aos respectivos livros de registro e, portanto, não conseguem verificar automaticamente o proprietário dos ativos e se estes podem ser transferidos. Em vez disso, muitos intermediários atuam como garantidores de ativos, já que o registro de transações atravessa organizações e os livros de registros são atualizados individualmente.

Existem diversas abordagens que estão sendo adotadas para o uso de *blockchain*. Uma abordagem de baixo risco é usar a *blockchain* internamente como um banco de dados para gerenciar ativos físicos e digitais, registrar transações internas, verificar identidades e outras aplicações. Essa pode ser uma solução especialmente útil para as empresas que têm dificuldade de conciliar múltiplos bancos de dados internos. O teste de aplicações de uso único pode ajudar as empresas a desenvolver as habilidades de que precisam para aplicações ainda mais avançadas.

Conforme revisão da literatura existente realizada por Lindman, Rossi e Tuunainen (2017), foi observado que há lacunas de pesquisa claras no estado da arte relacionado a esses sistemas descentralizados baseados na tecnologia *blockchain*. Segundo os autores, existe um apelo urgente para pesquisas sobre perspectivas teóricas e práticas sobre a tecnologia *blockchain*. Muitas questões críticas estão relacionadas à natureza inerentemente descentralizada desses sistemas de pagamento e plataformas.

O maior foco dos estudos está relacionado ao uso de criptomoedas, como o *bitcoin*, por exemplo. Ainda, os poucos estudos existentes procuram focar em uma característica ou uso específico destas tecnologias. Também é importante salientar que a grande massa de informações sobre o fenômeno está presente nas notícias publicadas na Internet e não em artigos científicos. O estudo proposto pretende realizar uma pesquisa geral nas notícias, mas com foco em encontrar conhecimento relacionado ao fenômeno *blockchain*.

É importante identificar quais tópicos já foram estudados e abordados em relação aos efeitos trazidos para a estratégia pelo fenômeno *blockchain* e quais são, atualmente, os maiores desafios e limitações que precisam de mais estudos. Para resolver essas questões, propõe-se um estudo de mapeamento para identificar como esse fenômeno está impactando a estratégia. Assim sendo, pretende-se

compreender: **De que forma a estratégia tecnológica tem sido afetada pelo fenômeno *blockchain*?**

1.2 OBJETIVOS

1.2.1 Objetivo Geral

Compreender de que forma o fenômeno *blockchain* tem afetado a estratégia tecnológica.

1.2.2 Objetivos Específicos

- a) Compreender quais impactos foram gerados pelas implementações de *blockchain* na estratégia tecnológica nos diversos setores;
- b) Mapear como o foco estratégico tecnológico das empresas está impulsionando o desenvolvimento desta tecnologia;
- c) Identificar o estágio de maturidade desta tecnologia.

1.3 JUSTIFICATIVA

Do ponto de vista da gestão, o fenômeno *blockchain* traz uma série de incertezas, por parte das organizações, sobre as suas reais implicações nas estratégias pós-plataforma. Contudo, muitas organizações estão realizando altos investimentos em experimentos de aplicações destas tecnologias. Um estudo recente da empresa britânica *Juniper Research* (HOLDEN; MOAR, 2017) anunciou que 57% das maiores empresas do mundo estão olhando para a tecnologia de *blockchain* como uma solução potencial integrada para uma série de problemas.

Para o autor do presente trabalho, esse assunto é muito instigante, pois propõe uma mudança disruptiva ao modelo das plataformas. Muitos especialistas da TIC dizem que este fenômeno é a maior evolução tecnológica desde a criação da Internet e que trará benefícios ainda maiores que os oferecidos pela Internet.

Para a teoria, faz-se necessário um mapeamento de como esse fenômeno vem moldando a estratégia tecnológica, mas também como ele vem sendo explorado no campo teórico. É importante ressaltar que existem diversos estudos relacionados ao

tema, mas a maioria relacionada ao uso de criptomoedas. Neste sentido, não foram encontrados estudos que façam uma pesquisa como a que se propõe neste trabalho. Nesse estudo de mapeamento, foram pesquisadas notícias vinculadas ao tema nos últimos 3 anos. O mapa produzido pela pesquisa atual sobre *blockchain* ajudará outros pesquisadores e profissionais na identificação de possíveis áreas de pesquisa e demais questões para futuros estudos.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo é o objeto da análise de múltiplas teorias e traz a literatura que dá o corpo para o tema estudado, bem como reúne o conhecimento necessário para a obtenção dos objetivos propostos nesse trabalho.

2.1 BLOCKCHAIN

A Internet está começando uma nova fase de descentralização. De acordo com Drescher (2017), após mais de vinte anos de pesquisa científica, houve grandes avanços nos campos da criptografia e das redes computacionais descentralizadas, resultando no surgimento de uma nova tecnologia profunda, conhecida como a *blockchain*, que tem o potencial de mudar fundamentalmente o modo como a sociedade opera. Drescher (2017) afirma que a *blockchain* é um banco de dados distribuído, compartilhado e criptografado que serve como um repositório público, irreversível e incorruptível de informações. Ele permite que pessoas não relacionadas possam chegar a um consenso sobre a ocorrência de uma determinada transação ou evento sem a necessidade de uma autoridade de controle.

De acordo com Tapscott e Tapscott (2016), a *blockchain*, que segue os princípios empregados no conceito criado para o *bitcoin*, é distribuída e funciona em computadores fornecidos por voluntários em todo o mundo, ou seja, não há um banco de dados central que possa ser comprometido. A *blockchain* é pública, pois qualquer pessoa pode visualizá-la a qualquer momento porque ela reside na rede, não dentro de uma única instituição encarregada de fazer auditorias de transações e manter registros. A *blockchain* é criptografada, pois usa criptografia pesada envolvendo chaves públicas e privadas, como o sistema de duas chaves, para acessar uma caixa de segurança e manter a segurança virtual.

Conforme Drescher (2017), alcançar a integridade e a confiança nas melhores condições é fácil. O verdadeiro desafio é alcançar a integridade e a confiança em um sistema distribuído onde os membros desse sistema não confiam uns nos outros. Este é o problema que a *blockchain* deve resolver. O principal problema a ser resolvido pela *blockchain* é alcançar e manter a integridade em um sistema distribuído de forma pura, que consiste em um número desconhecido de pares com confiabilidade desconhecida. Esse problema não é novo. Na verdade, é um problema bem

conhecido e amplamente discutido na ciência da computação. Ao utilizar uma metáfora das forças armadas, o problema é amplamente considerado como o problema dos generais bizantinos (LAMPORT; SHOSTAK; PEASE, 1982).

Segundo Lindman, Rossi e Tuunainen (2017), a tecnologia *blockchain* e as tecnologias de banco de dados distribuídas subjacentes são os principais fatores tecnológicos dos desenvolvimentos recentes em sistemas de transações distribuídas. Os instrumentos financeiros, tais como pagamentos, registros comerciais e contratos inteligentes, podem ser construídos em tecnologia *blockchain*, o que evita comportamentos e repercussões adversas, como gastos duplos, falsificações e falsas disputas (BARBER et al., 2012). Além disso, a tecnologia pode ser usada para registros legais e públicos, como títulos, certidões de nascimento, votação ou registros judiciais. Ela também pode ser usada para a criação de *smart property*³, caso em que o *blockchain* se torna um inventário, realiza o rastreamento e serve de mecanismo *buysell*⁴ para ativos duráveis, como diamantes ou carros. Outra forma de uso seria para rastrear a criação de produtos para negócios socialmente responsáveis. De acordo com Mainelli e Smith (2015), *blockchain* pode ser utilizado como um mecanismo transacional para os serviços de compartilhamento de economia, pois naturalmente resolve a gravação confiável de atividades em grande escala. A importância desse mecanismo transacional aumenta com um mundo cada vez mais conectado, onde uma quantidade crescente de coisas físicas se torna programável e conecta-se à Internet.

Para Tapscott e Tapscott (2016), grandes bancos e alguns governos estão implementando *blockchains* como livros contábeis distribuídos para revolucionar a forma como as informações são armazenadas e as transações ocorrem. Seus objetivos são alcançar velocidade e segurança, ter menor custo, diminuir os erros e eliminar pontos centrais de ataque e falha.

Em resumo, a força da tecnologia *blockchain* deriva diretamente de três fatores e da maneira como eles interagem entre si: (1) a natureza distribuída do livro razão gera transparência e sincronização; (2) o protocolo de consenso anula a necessidade

³ São as propriedades nas quais as informações de posse são controladas por contratos da *blockchain* e que também ficam registradas na mesma (ACHESON, 2015).

⁴ É um acordo juridicamente vinculativo entre coproprietários de uma empresa que rege a situação se um coproprietário morrer ou for forçado a deixar a empresa ou optar por deixar a empresa.

de confiança; e (3) a forma como os dados são gravados, armazenados e conectados produz imutabilidade e rastreabilidade (PISA; JUDEN, 2017).

De acordo com Kabashkin (2013), o crescimento futuro e a adoção de *blockchain* dependem da construção de um forte ecossistema. A rede de ecossistemas *blockchain* é uma plataforma global e transversal para a colaboração, a rede, a explicação e o avanço das metodologias e soluções *blockchain* para os setores público e privado.

Apenas a melhor compreensão desses riscos e oportunidades levará a serviços melhores, mais confiáveis e mais eficientes para os cidadãos, os consumidores e a variedade de organizações com interesse no desenvolvimento de tecnologias de *blockchain*. Para entender a lógica da operação e as oportunidades e riscos resultantes dos serviços e plataforma da tecnologia *blockchain*, Lindman, Rossi e Tuunainen (2017) propõem focar em fatores organizacionais, ambientais e tecnológicos.

Os fatores organizacionais incluem a organização e gestão do serviço focal ou da plataforma, incluindo questões como estrutura financeira (estrutura de custos e potencial de lucro do serviço) e as estratégias de preços relacionadas ou, em outras palavras, o modelo de negócios subjacente (HEDMAN; KALLING, 2003). Os fatores ambientais, por sua vez, incluem a demanda, ou seja, os usuários, bem como a concorrência, tanto em relação aos concorrentes diretos que oferecem serviços similares quanto às empresas que oferecem serviços de substituição. Já os fatores de tecnologia, referem-se aos detalhes das tecnologias de informação utilizadas, com especial interesse nas questões de *design* relacionadas ao desenvolvimento desses sistemas.

Em essência, os sistemas e plataformas de pagamento distribuídos e outros serviços habilitados pela tecnologia *blockchain* podem ser conceitualizados como inovações de serviços. Autores apontam que essas inovações de serviços melhoram a produtividade dos serviços ou desenvolvem novos modelos de serviços (AVLONITIS; PAPASTATHOPOULOU; GOUNARIS, 2001). Assim, conforme Bishop et al. (2008), pode-se entender as inovações de serviços como a combinação de inovação tecnológica, inovação em modelos de negócios, inovação social-organizacional e inovação de demanda com o objetivo de melhorar os sistemas de serviço existentes (inovação incremental), criar novas proposições de valor (ofertas) ou criar novos sistemas de serviço (inovação radical).

Segundo Lindman, Rossi e Tuunainen (2017), as inovações de serviços são, em geral, e no setor financeiro em particular, muitas vezes, baseadas em tecnologia, seja na introdução de uma nova tecnologia ou no uso diferente da tecnologia existente. As moedas digitais descentralizadas, como o *bitcoin*, podem ser vistas como extensão de linha de serviço nova ou comercial, inovação tecnológica ou demanda de inovação, ou algum outro tipo de inovação no serviço, dependendo da perspectiva adotada. Independentemente da perspectiva, as oportunidades, bem como os riscos, são distintamente diferentes, como por exemplo, da perspectiva dos bancos e dos legisladores para os consumidores individuais ou comerciantes.

Com a tecnologia *blockchain*, os problemas organizacionais mais críticos por enquanto estão relacionados às plataformas de pagamento digital e, mais detalhadamente, os efeitos de rede aos quais esses tipos de plataformas estão sujeitos. Outro conjunto de questões organizacionais chave pode ser derivado dos novos modelos de negócios habilitados e facilitados pela tecnologia *blockchain*.

De forma complementar, segundo Economides e Katsamakos (2006), o sucesso das redes e serviços mediados por plataforma depende muito do tamanho da rede de usuários. Os efeitos de rede são os efeitos positivos ou negativos que as ações de um usuário têm na avaliação de outros usuários da rede (MILGROM; ROBERTS, 1992). Em outras palavras, de acordo com Katz e Shapiro (1994), o valor da associação a um usuário é afetado, positivamente, por outro usuário se juntando e ampliando a rede.

No desenvolvimento dos pagamentos digitais, a confiança permaneceu um enigma (SHAW, 2014). A literatura de pagamento móvel argumenta sobre a necessidade de um gerenciador de serviços confiável que administre autenticação, autorização e liquidação de contas (ONDRUS; PIGNEUR, 2009), especialmente na presença de efeitos de rede direta e indireta, conforme descrito por Au e Kauffman (2008). As plataformas de transações distribuídas estão respondendo a esse desafio por uma descentralização radical baseada em redes *peer-to-peer*. Estas são chamadas de moedas digitais descentralizadas ou criptomoedas e sistemas de consenso descentralizado (GLASER; BEZZENBERGER, 2015).

O fenômeno *blockchain* vem sendo amplamente discutido, e acredita-se que ele esteja sendo usado como parte da estratégia tecnológica. Ao combinar essas tecnologias com mudanças estruturais, como a cultural e a regulatória, acredita-se que se iniciará uma transformação no modo como as empresas operam. Assim, dando

prosseguimento a temática, aprofunda-se o conceito da estratégia tecnológica afim de trazer a literatura acerca do assunto.

2.2 ESTRATÉGIA TECNOLÓGICA

O mercado vem impondo novos desafios às organizações. Os produtos têm reduzido seus ciclos de vida de forma contínua, exigindo das empresas o reconhecimento e análise das novas oportunidades tecnológicas. Além disso, novas tecnologias são constantemente criadas para as mais diferentes aplicações (CHERUBIN, 2000).

A partir desses desafios, a estratégia tecnológica surge para auxiliar no processo de decisão quanto às tecnologias a serem adotadas, sempre se considerando as tendências tecnológicas e as necessidades do mercado. Além disso, Cherubin (2000) destaca que, à medida que grandes quantidades de tecnologias são disponibilizadas, faz-se necessário um processo de seleção das tecnologias mais relevantes aos propósitos da empresa. Seguindo nessa linha, Motta (1995, p. 13) destaca que

na era da competitividade global, o grande desafio das empresas está centrado na capacidade de busca de novas tecnologias, novos mercados e novos métodos de gerenciamento ou do redesenho dos processos de negócios e de integração total das cadeias de valor da empresa, clientes e fornecedores.

É importante observar que a essência da estratégia tecnológica de uma organização reside nos conhecimentos e habilidades que ela dispõe, ou pretende dispor, e não nos produtos que ela tem ou nos mercados em que ela atua. Neste sentido, segundo Ford (1988), a estratégia tecnológica consiste em políticas, planos e procedimentos para adquirir conhecimentos e habilidades, administrá-los dentro da organização e explorá-los para obter o maior retorno possível.

Desta forma, é conveniente que um grupo de organizações esteja mais disposto para pesquisar e explorar com mais intensidade a tecnologia, e outros apenas utilizem destes novos conceitos. Deste modo, analisar em quais tecnologias as empresas devem investir seus recursos, não apenas individualmente, mas como parte de seu portfólio de tecnologias, é uma decisão estratégica (CLARKE et al., 1995). Além disso, essas empresas podem decidir quanto ao meio mais apropriado para a aquisição, podendo variar desde menores ou maiores integrações, como utilizar sua própria equipe de Pesquisa e Desenvolvimento (P&D), estabelecer

contratos de pesquisa em conjunto ou totalmente externos e, ainda, utilizarem-se dos licenciamentos (CLARKE et al., 1995).

De acordo com Porter (1992), a estratégia tecnológica é o meio utilizado por uma empresa para o desenvolvimento e o uso de tecnologias. No centro desta estratégia tem-se o tipo de vantagem competitiva que a empresa busca. As tecnologias a serem desenvolvidas pela organização são aquelas que trazem maior contribuição para a estratégia geral da empresa. Nessa ótica, a estratégia tecnológica deve abordar três questões centrais: (1) as tecnologias a serem desenvolvidas; (2) se a liderança tecnológica deve ser buscada nestas tecnologias; e (3) o papel do licenciamento de tecnologia (PORTER, 1992).

Clarke et al. (1995) afirmam que uma estratégia tecnológica envolve outros elementos além da estratégia de P&D, a política de desenvolvimento de produto ou a estratégia de operações. Ela deve envolver a aquisição de tecnologia, a sua exploração e o seu gerenciamento. Ford (1988) ressalta que a importância desta estratégia está na análise das bases tecnológicas de suas estruturas de produção e produtos.

A estratégia tecnológica apoia a estratégia corporativa de uma empresa, mostrando como os objetivos gerais da empresa podem ser alcançados do ponto de vista tecnológico (BROCKHOFF; PEARSON, 1992). Em particular, a estratégia de tecnologia corporativa reflete o nível de agressividade tecnológica de uma empresa, que constitui uma dimensão importante desta estratégia, sendo abordada em múltiplos estudos. Miles et al. (1978) basearam-se em ideias sobre a agressividade tecnológica na diferenciação entre estratégias tecnológicas proativas ou progressivas e reativas ou conservadoras. Já Brockhoff e Pearson (1992) diferenciaram altos e baixos graus de agressividade tecnológica nas estratégias de inovação.

Miles et al. (1978) criaram uma tipologia de quatro classificações que analisa as organizações quanto a sua disposição para mudança. A dimensão principal que traz diferença para estas classificações é justamente a velocidade com a qual a empresa responde a uma mudança no ambiente, com a respectiva alteração dos seus produtos e mercados. As quatro classificações são: prospectores, defensores, analisadores e reativos.

Conforme Griffin e Page (1996), os prospectores acreditam que o valor deve ser priorizado para novos produtos, mercados e tecnologias, embora nem todos os esforços se tornem lucrativos. Eles respondem de forma ágil para sinais iniciais de

mudanças do ambiente preocupando-se com as áreas de oportunidade que podem ser exploradas. Em relação aos defensores, Griffin e Page (1996) explicam que o seu foco é o de encontrar e manter um nicho seguro em um produto ou serviço relativamente estável. Eles protegem o seu domínio oferecendo alta qualidade, serviços superiores e preços baixos. Estas empresas ignoram as mudanças na indústria que não tenham uma influência direta em suas operações correntes. Quanto aos analisadores, raramente são os primeiros do mercado a lançar novos produtos. Entretanto, através do atento monitoramento das ações de seus concorrentes, eles podem frequentemente ser um seguidor rápido, trazendo um processo mais eficiente, uma redução de custos ou um produto inovador ao mercado de forma muito rápida. Já os reativos, não são tão agressivos para manter os seus produtos estabelecidos e os mercados onde estão inseridos quanto os seus competidores. Eles respondem somente quando são forçados por pressões geradas por mudanças ambientais (MILES et al., 1978).

De acordo com Griffin e Page (1996), uma organização prospectora buscará desenvolver pioneiramente um novo produto ou processo, mesmo que os lucros por esta atividade ocorram por um período pequeno de tempo. As organizações defensoras preocupam-se em diversificar, aumentando a penetração de mercado e defendendo seu nicho com produtos que apresentam maior qualidade ou menor preço. As analisadoras raramente apresentam um novo conceito ao mercado, geralmente são as seguidoras rápidas. Estas empresas buscarão desenvolver uma maior eficiência de custo ou de desempenho dos produtos no mercado. As reativas, no entanto, respondem somente quando forçadas pelos competidores (MILES et al., 1978).

Com isso, percebe-se qual a disposição para a inovação que cada organização pretende e mostra ao mercado, sendo possível discutir sobre algumas vantagens oriundas da abordagem escolhida pelas organizações. Por exemplo, no caso da escolha por uma estratégia prospectora, a organização fará altos investimentos em pesquisa e desenvolvimento. Porém, aproveita-se do potencial de explorar economicamente este novo conceito, o que representa um diferencial tecnológico para aquela organização. Ao optar por uma estratégia analista ou defensora, embora exista uma diferença quanto à intensidade do esforço tecnológico entre estas duas abordagens, as organizações procuram oferecer tecnologias incrementais. Estas tecnologias requerem um menor esforço de P&D e contribuem para reduzir os custos

desta atividade e, conseqüentemente, também reduzem os seus riscos inerentes. Entretanto, não são identificadas vantagens advindas da escolha por uma estratégia reativa, pois a inexistência de ações dificulta esta análise (GRIFFIN; PAGE, 1996).

Em empresas com estratégias de tecnologia agressivas, o desenvolvimento do conhecimento tecnológico proprietário constitui uma dimensão importante da estratégia corporativa (GRANSTRAND, 1998). Assim, a propriedade exclusiva das novas tecnologias é essencial para ganhar e manter uma vantagem competitiva claramente baseada em tecnologia (BROCKHOFF; PEARSON, 1992; MILES et al., 1978). Muitas dessas empresas se concentram em desenvolver constantemente novas tecnologias que são superiores às tecnologias de seus concorrentes, caracterizadas por uma orientação tecnológica em seus processos de inovação (GATIGNON; XUEREB, 1997). Segundo Deeds (2001), os efeitos positivos da agressividade tecnológica são particularmente fortes nas configurações de alta tecnologia, onde novos produtos exigem um alto grau de novidade para acompanhar a concorrência. De acordo com Teece (2006), os ambientes de alta tecnologia tendem a ser mais dinâmicos e as empresas devem ser inovadoras nessas configurações dinâmicas para evitar ficar atrás de seus concorrentes. Além disso, as estratégias de tecnologia agressivas requerem um nível mínimo de *expertise* e entrada de recursos, ou seja, despesas de P&D, para serem efetivas (GARCIA-VEGA, 2006). Conseqüentemente, os efeitos positivos da agressividade tecnológica provavelmente serão mais fortes em ambientes de alta tecnologia.

Walsh e Linton (2011) criaram um modelo chamado *Strategy-Technology Firm Fit Audit*. Sua intenção de *design* é utilizar esse modelo para cenários de inovação aberta, empresas empreendedoras e empresas já estabelecidas. A base elementar deste modelo é que a experiência passada leva a uma compreensão da combinação atual de competências e capacidades de uma empresa. Além disso, o estado atual da empresa dá uma visão de sua capacidade de abraçar as oportunidades futuras. Se constitui como um processo interno abrangente e efetivo para a revisão de oportunidades, com base em sinergia com as capacidades gerenciais da empresa e as competências tecnológicas.

De acordo com estudo realizado por Arasti, Khaleghi e Noori (2017), quase todas as pesquisas defendem a posição de estratégia tecnológica na hierarquia de estratégias de uma empresa como uma estratégia funcional no nível da unidade de negócios. Com base nas pesquisas realizadas, os autores afirmam que “com base

nos nossos melhores conhecimentos, não há estrutura ou modelo que reconheça explicitamente a estratégia tecnológica no nível corporativo” (ARASTI; KHALEGHI; NOORI, 2017, p. 2).

A tecnologia tem ação transformadora, tanto nos aspectos internos à organização quanto nos aspectos relacionados ao ambiente em que essa empresa atua. Por essa razão, esta pesquisa assume a importância dos estudos relativos à estratégia tecnológica das empresas, assim como do fenômeno tecnológico *blockchain*, devido à crescente busca de iniciativas relacionadas a esse fenômeno nas organizações, nos últimos anos, em todo o mundo. No entanto, dispor de tecnologias avançadas ou transformadoras não traz grandes contribuições à organização, se estas não puderem ser assimiladas por conhecimentos necessários para o seu efetivo desenvolvimento e uso. Pode-se dizer que compreender os aspectos relacionados ao conhecimento é tão relevante quanto à própria tecnologia desenvolvida ou adquirida pela organização.

Cada vez mais, as organizações estão focadas em P&D, pois a todo momento são disponibilizadas novas tecnologias que podem resolver problemas, aumentar a eficiência ou reduzir custos operacionais dos seus produtos e serviços. Porém, é importante salientar que esses estudos precisam trazer resultados, visto que os investimentos em P&D são expressivos.

Ao longo dos últimos anos, onde a tecnologia da informação evoluiu muito rápido graças a expansão da Internet, muitas empresas foram surpreendidas por tecnologias que mudaram o ambiente e os mercados de forma abrupta, bem como por, em muitos casos, *startups* que surgiram como novos entrantes em mercados antes vistos como dominados por grandes organizações. Segundo essa perspectiva, faz-se importante, neste estudo, compreender os aspectos relacionados à influência que o fenômeno *blockchain* trouxe para a estratégia tecnológica.

Após as considerações feitas na base teórico-empírica desenvolvida, aborda-se, na parte seguinte, a proposta metodológica que busca orientar a investigação do problema de pesquisa proposto, levando-se em conta as variáveis, o delineamento da pesquisa e os procedimentos para a coleta e o tratamento dos dados.

3 MÉTODO

Este capítulo aborda o método e os procedimentos metodológicos utilizados para alcançar os objetivos propostos neste trabalho. Ele está dividido em delineamento da pesquisa, procedimentos de coleta de dados e procedimentos de análise dos dados.

3.1 DELINEAMENTO DA PESQUISA

As empresas de tecnologia de hoje dependem, em grande parte, de insumos numéricos e categóricos para ganhos de informação, algoritmos de aprendizado de máquina ou otimização operacional. Desta maneira, faz sentido para uma organização estudar apenas informações estruturadas. O texto representa uma entrada inexplorada que pode aumentar ainda mais a vantagem competitiva.

Neste trabalho, foi realizada uma pesquisa exploratória com abordagem quantitativa utilizando técnicas de Descoberta de Conhecimento em Textos (DCT) com o uso de ferramentas de mineração de textos. É importante mencionar que a escolha deste processo se dá em face a grande existência de dados não estruturados acerca do tema veiculados através de notícias em sites na *Web*, bem como pela dificuldade de leitura e análise manual de todas essas notícias que, através deste processo, pode ser automatizada. Através do estudo teórico acerca do assunto e dos achados referentes às análises realizadas através das técnicas de DCT, foram criadas proposições relacionadas aos conhecimentos encontrados e aos objetivos deste trabalho.

3.1.1 Descoberta de Conhecimento em Textos

O processo capaz de gerar conhecimento a partir de dados estruturados é nomeado Descoberta de Conhecimento em Bases de Dados (DCBD). Esse processo combina diversas áreas da descoberta do conhecimento, tais como Aprendizagem de Máquina, Reconhecimento de Padrões, Estatística e Inteligência Artificial, com o objetivo de extrair, de forma automática, informação útil de bases de dados (SILVA et al., 2016). Já a DCT lida com dados não-estruturados. Muitas pesquisas têm sido direcionadas à DCT por trabalhar com textos, considerada a forma mais natural de

armazenamento de informação (TAN, 1999). Para o desenvolvimento deste trabalho, foi utilizada a técnica de Descoberta de Conhecimento em Texto.

A necessidade de obter-se conhecimento através de textos com dados não estruturados é crescente, visto a quantidade exponencial de notícias e informações que são publicadas diariamente na Internet. Conforme Tan (1999), estatisticamente, 80% das informações de uma organização estão disponíveis de forma textual não estruturada. No entanto, encontrar conhecimento em meio a todas essas informações não estruturadas é uma tarefa complexa. O termo DCT foi proposto, inicialmente, por Feldman e Dagan (1995) para descrever o processo de encontrar informações importantes em grandes coleções de textos (artigos, revistas e jornais, páginas *Web*, etc.). Atualmente, também são utilizados como sinônimo de DCT os termos *Text Mining* ou *Text Data Mining* (TAN, 1999).

Conforme Tan (1999), o DCT pode ser definido como o processo para extração de padrões ou conhecimentos interessantes e não-triviais a partir de documentos textuais, ou seja, essa área possibilita a transformação de informações não estruturadas em informações estruturadas. Pode-se dizer que descoberta de conhecimento em textos é uma evolução natural da recuperação de informações, já que os sistemas de recuperação de informações passaram a adotar algumas técnicas de análise de informações, muitas das quais provenientes da área de descoberta de conhecimento em bases de dados. Assim, ao invés do pesquisador ter que identificar se documentos retornados são realmente relevantes, o próprio sistema faz essa análise e retorna as informações de forma condensada e resumida (LEAL, 2010).

Ainda segundo Leal (2010), a descoberta de conhecimento em textos pode ser entendida como a aplicação de técnicas de descoberta de conhecimento sobre dados extraídos de textos. Entretanto, cabe salientar que descoberta de conhecimento em textos não inclui somente a aplicação das técnicas tradicionais de descoberta de conhecimento em dados, mas também qualquer técnica nova ou antiga que possa ser aplicada no sentido de encontrar conhecimento em qualquer tipo de texto. Com isso, muitos métodos foram adaptados ou criados para suportar esse tipo de informação semiestruturada ou sem estrutura, que é o texto.

De acordo com Ebecken, Lopes e Costa (2003), existem dois tipos de abordagens para análise de dados textuais na área de mineração de textos: a Análise Semântica, baseada na funcionalidade dos termos encontrados nos textos, e a

Análise Estatística, baseada na frequência dos termos encontrados nos textos. Estas abordagens podem ser utilizadas separadamente ou em conjunto.

A Análise Semântica emprega técnicas que avaliam a sequência dos termos no contexto dos textos, no sentido de identificar qual a sua função. Ela é fundamentada em técnicas de Processamento de Linguagem Natural (PLN). Sua utilização justifica-se, principalmente, pela melhoria da qualidade dos resultados do processo de mineração de textos, especialmente se for incrementado por Processamento Linguístico (LUÍS; ROSA, 1998).

De acordo com Aranha e Passos (2006), o PLN é uma técnica chave para a DCT. Utilizando conhecimentos da área de linguística, o PLN permite aproveitar ao máximo o conteúdo do texto, extraindo entidades, seus relacionamentos, detectando sinônimos, corrigindo palavras escritas de forma errada e, ainda, as desambiguizando. Participam normalmente na parte do pré-processamento dos dados, transformando-os em números.

Segundo Miranda (2009), na Análise Estatística, os termos são valorados, basicamente, pela sua frequência de aparição na massa de dados, não importando a contextualização deste, como em que parágrafo está inserido, que termos o antecedem ou quais estão diretamente relacionados. Neste tipo de análise, a importância de um termo é dada pelo número de vezes que este aparece no texto. Basicamente, seu processo envolve aprendizado estatístico a partir de dados, que normalmente inclui as etapas de codificação dos dados, estimativa dos dados e modelos de representação de documentos.

O processo da DCT como um todo, conforme proposto por Aranha (2007), sugere um modelo completo para aquisição de conhecimento a partir de textos, com um especial aprofundamento na etapa de pré-processamento. A Coleta é a etapa inicial e tem como função formar a base de textos que são relevantes para alcançar o objetivo almejado, isto é, determinar e selecionar o universo de atuação das técnicas de mineração de texto.

A etapa de Pré-processamento é executada imediatamente após a Coleta e tem como objetivo prover alguma formatação e representação da massa textual. É bastante onerosa, com a aplicação de diversos algoritmos que consomem boa parte do tempo do processo de extração de conhecimento. Já a Indexação é o processo que organiza todos os termos adquiridos a partir das fontes de dados, facilitando o

seu acesso e recuperação. Uma boa estrutura de índices garante rapidez e agilidade ao processo, tal como funciona o índice de um livro.

Após terem sido obtidas uma estrutura para os dados e uma forma de prover rápido acesso, a etapa de Mineração propriamente dita é responsável pelo desenvolvimento de cálculos, inferências e algoritmos e que tem como objetivo a extração de conhecimento, descoberta de padrões e comportamentos que possam surpreender de acordo com o objetivo e requisitos do usuário e/ou domínio da aplicação. Finalmente, a Análise é a última etapa e deve ser executada por pessoas que, normalmente, estão interessadas no conhecimento extraído e que devem tomar algum tipo de decisão apoiada no processo de mineração de texto.

Segundo Loh (2001), apesar da possibilidade do processo de descoberta ser iniciado sem que o usuário defina proposições (abordagem proativa), a intervenção humana ainda se faz necessária. Para levantar ou validar proposições, é necessário interpretar os resultados sob o contexto do domínio. Somente pessoas ligadas ao domínio podem filtrar os resultados para extrair conhecimento interessante. Em alguns casos, o conhecimento descoberto pode ser útil, mas não novo. Em outros casos, ele pode ser novo, mas não ser útil. Somente as pessoas do domínio podem avaliar o que é útil e novo e isto sim se torna conhecimento interessante. Neste caso, a intervenção humana serve para filtrar o que é interessante (útil e novo).

Desta forma, o conhecimento prévio (*background knowledge*) sobre o domínio também é importante no processo de descoberta de conhecimento em textos. Este conhecimento pode ser usado pelo usuário para limitar o espaço de pesquisa, para analisar ou para definir proposições (LOH, 2001). Também pode-se usar conhecimento prévio na definição dos conceitos ou na interpretação dos resultados. Neste caso, ele é útil para eliminar ambiguidades. Feldman e Hirsh (1997) aplicam técnicas de descoberta sobre palavras, permitindo que o usuário intervenha no processo, fazendo uso de seus conhecimentos prévios sobre o domínio ou assunto. Isto acelera o processo e permite filtrar os resultados de acordo com o interesse do usuário. Tendo como suporte o processo de DCT, este trabalho seguiu os procedimentos de coleta descritos a seguir.

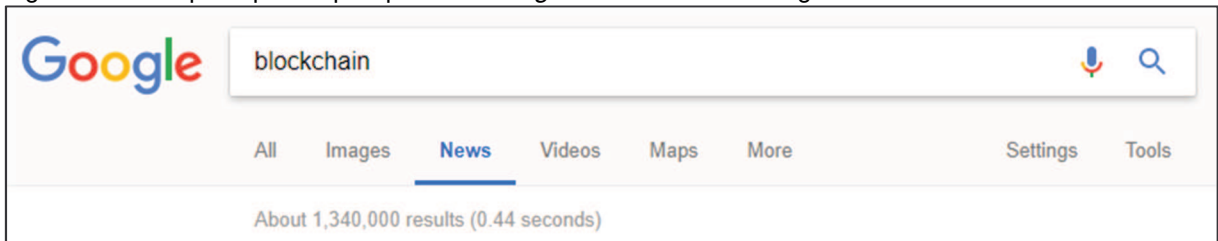
3.2 PROCEDIMENTOS DE COLETA DE DADOS

A primeira etapa do processo de DCT é a Coleta dos textos que serão utilizados na pesquisa. Quanto à origem dos textos, os mesmos podem ser obtidos das mais diversas fontes, mas, em geral, são três os principais ambientes de localização: pastas de arquivos encontradas no disco rígido dos usuários, tabelas de diversos bancos de dados e *Web* (SOARES, 2008).

Este trabalho utilizou uma técnica de coleta de dados na *Web* dividida em duas fases. A primeira fase consistiu na pesquisa, através de ferramenta de busca de notícias, de notícias acerca do tema *blockchain* e da extração de suas *Uniform Resource Locator* (URLs). A segunda fase consistiu na extração dos textos das notícias utilizando um *software* de *Web Crawler*⁵, seguido de uma verificação geral dos dados extraídos.

Durante a primeira etapa, foi utilizada a ferramenta de pesquisa *Google News*, onde realizou-se uma pesquisa por notícias no idioma Inglês. Essas notícias deveriam apresentar o termo *blockchain* em qualquer parte do seu conteúdo, conforme mostrado na Figura 1.

Figura 1 – Tela principal da pesquisa no *Google News* no idioma Inglês

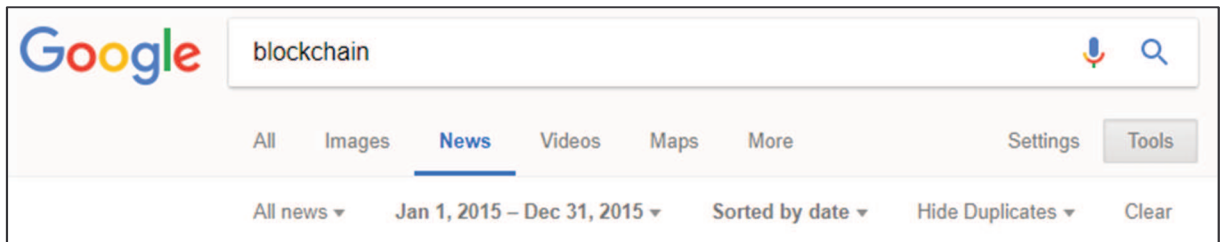


Fonte: Google (2018).

Foi adicionada uma regra de pesquisa para que fossem listadas somente as notícias publicadas no período de 1º de janeiro de 2015 a 31 de dezembro de 2017. Os dados de cada ano foram coletados separadamente. Além disso, foram configuradas regras de pesquisa visando a remoção de notícias duplicadas, bem como para que a pesquisa mostrasse 100 resultados por página, de forma a deixar o processo de extração das URLs mais simples. A Figura 2 mostra a configuração usada para remoção de notícias duplicada durante a pesquisa de notícias do ano de 2015.

⁵ É um rastreador da *Web*, às vezes chamado de *spider*. É um *bot* da Internet que navega sistematicamente na *World Wide Web*, normalmente para fins de indexação da *Web*.

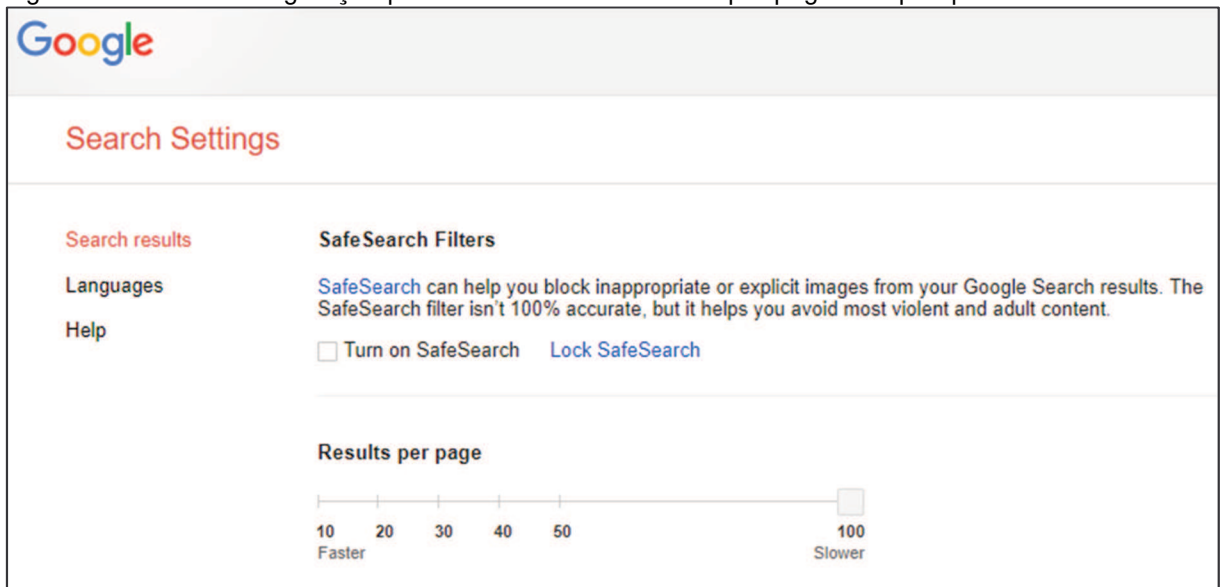
Figura 2 – Configuração de range de data da coleta e remoção de notícias duplicadas, referente ao ano de 2015



Fonte: Google (2018).

A Figura 3 apresenta a tela de configuração para mostrar 100 resultados por página durante a pesquisa.

Figura 3 – Tela de configuração para mostrar 100 resultados por página de pesquisa



Fonte: Google (2018).

O período inicial desta pesquisa foi escolhido com base em análise realizada através da ferramenta *Google Trends*, onde detectou-se que o tema começou a ter um aumento significativo de pesquisas na *Web* no ano de 2015. É importante mencionar que o algoritmo de pesquisa do Google traz, no seu resultado de pesquisa, as notícias com maior relevância encontradas na *Web* e está limitado a apresentar, no máximo, 1.000 resultados por pesquisa. A Figura 4 mostra a evolução da pesquisa do termo *blockchain* entre 2015 e 2017 no *Google Trends*.

Figura 4 – Evolução da pesquisa pelo termo *blockchain* entre 2015 e 2017 (Google Trends)



Fonte: Google Trends (2018).

Após a realização da pesquisa de URLs referentes a cada ano da pesquisa, foi utilizado o *software* Google *SERPs Extractor* (AINSWORTH, 2015), desenvolvido em linguagem de programação *JavaScript*⁶, para listar todas as URLs de notícias. Isso permitiu a cópia destes endereços em uma planilha Excel. O código deste *JavaScript* encontra-se em apêndice neste trabalho (ANEXO A – Código *JavaScript* do Google *SERPs Extractor*). Após copiar as URLs para a planilha Excel, os dados foram organizados em ordem alfabética de modo que as URLs dos sites ficassem organizadas de forma unificada. Nesta etapa, foram mapeadas um total de 2.768 URLs de notícias. Após a criação e organização da lista de URLs, iniciou-se a segunda etapa do processo de coleta com a extração da notícia propriamente dita através do *software* de *Web Crawler*. Imagens parciais das planilhas de URLs geradas para esta pesquisa estão disponíveis em apêndice neste trabalho, conforme dados de 2015 (APÊNDICE A – Planilha de URLs do ano de 2015), 2016 (APÊNDICE B – Planilha de URLs do ano de 2016) e 2017 (APÊNDICE C – Planilha de URLs do ano de 2017).

Na segunda etapa, foi utilizado o *software* *Octoparse* (OCTOPARSE, 2018), responsável pela extração dos títulos e textos das notícias referentes às URLs extraídas na primeira etapa da pesquisa. O processo de extração, apesar de realizado com o auxílio do *software*, teve uma grande quantidade de interações manuais devido à grande quantidade de notícias mapeadas, bem como pelos diferentes padrões

⁶ O *JavaScript*, geralmente abreviado como JS, é uma linguagem de programação interpretada de alto nível.

utilizados pelos portais de notícias, onde o formato dos campos de textos é diferente em cada portal. A extração se deu por afinidade de site publicador, ou seja, notícias publicadas em um mesmo portal eram extraídas em conjunto e armazenadas em um mesmo arquivo no formato .CSV. Na medida que as notícias foram extraídas com sucesso, às URLs referentes às mesmas foram marcadas, na planilha Excel, com o status de extraídas com sucesso. Do total de 2.768 notícias encontradas na primeira etapa, foram eliminadas 163 notícias devido à restrição de acesso, restando um total de 2.605 notícias extraídas com sucesso.

Após a extração das notícias, foi realizada uma verificação geral dos dados, pois o *software* de exportação trouxe diversos códigos de linguagem de programação HTML, que poluíram o *corpus* e precisaram ser removidos. Esta tarefa foi realizada através do desenvolvimento de um *Script* macro no Excel, criado pelo autor, e depois validada manualmente em cada documento da base de dados. Por fim, os arquivos gerados no formato .CSV foram convertidos para o formato .TXT de forma a facilitar a compatibilidade com os *softwares* de Mineração de Textos disponíveis no mercado. Com isso, a etapa de coleta de dados foi finalizada.

3.3 PROCEDIMENTOS DE ANÁLISE DE DADOS

Uma vez realizada a Coleta de dados, o próximo passo foi a preparação dos textos, utilizando-se o *software* R, para que os mesmos fossem manipulados pelos algoritmos de Mineração de Textos. Esta etapa foi denominada fase de Pré-processamento. Essa fase foi responsável por um conjunto de transformações realizadas sobre a coleção de textos com o objetivo de fazer com que esses passassem a ser estruturados em uma representação atributo-valor, para ser manipulada pelos métodos de extração de conhecimento.

Pré-processar textos é, por muitas vezes, o processo mais oneroso da metodologia de Mineração de Textos, uma vez que não existe uma única técnica que possa ser aplicada para a obtenção de uma representação satisfatória em todos os domínios, sendo necessária a realização de experimentos empíricos para se chegar à representação adequada (CARRILHO JÚNIOR, 2007). Nesta fase, foram realizadas diversas atividades para o tratamento dos textos. Estas atividades são citadas e explicadas a seguir.

O *software* adotado para analisar a descoberta de conhecimento em texto nesse estudo foi o R (ou *R-project*) versão 3.4.4, que pode ser obtido gratuitamente no *site* <<http://www.r-project.org>>. O R é um *software* livre e amplamente utilizado na área de estatística e *data mining*. Possui um conjunto de funções pré-definidas, mas outras funções podem ser instaladas através de pacotes que estão disponíveis para *download* no mesmo *site* onde encontra-se o *software*.

A base de dados foi formada por 466 arquivos, contendo um total de 2.605 notícias, todas no idioma inglês, sobre o tema *blockchain*. Para fazer uma análise dos termos mais frequentes nas notícias, as 2.605 notícias passaram por procedimentos de Pré-processamento.

O primeiro passo da operação de Pré-processamento foi a tokenização⁷ e sua execução teve, como finalidade, seccionar um documento textual em unidades mínimas, mas, que exprimam a mesma semântica original do texto. O termo *token* é utilizado para designar estas unidades, que em muitas vezes correspondem a somente uma palavra do texto. Porém, nem sempre estas unidades textuais podem ser consideradas palavras ou apresentam mais de uma palavra (SOARES, 2008). O processo de tokenização é auxiliado pelo fato das palavras serem separadas por caracteres de controle de arquivo ou de formatação, tais como espaços ou sinais de pontuação que, em alguns casos, podem ser considerados *tokens* delimitadores (FELDMAN; SANGER, 2007).

Uma vez realizado o processo de tokenização, o próximo passo foi a identificação das palavras que poderiam ser desconsideradas nos passos posteriores do processamento dos dados. Esta fase é a tentativa de limpeza de tudo que não constitui conhecimento nos textos. Em um documento, existem muitos *tokens* que não possuem nenhum valor semântico, sendo úteis apenas para o entendimento e compreensão geral do texto. Estes *tokens* são palavras classificadas como *stopwords*⁸ e fazem parte do que é chamado de *stoplist* de um sistema de Mineração de Textos (BASTOS, 2006). Uma *stoplist* bem elaborada permite a eliminação de muitos termos irrelevantes, tornando mais eficiente o resultado obtido pelo processo de Mineração de Textos. Normalmente, entre 40% e 50% do total de palavras de um

⁷ Do inglês, *tokenization*.

⁸ Palavras muito frequentes e com pouco significado (tais como artigos, preposições, algumas conjunções), geralmente são desconsideradas nas minerações de texto (LOH, 2008).

texto são removidas com uma *stoplist* (SILVA, 2007). Este trabalho utilizou a matriz referência de Lewis et al. (2004).

A Indexação foi a fase responsável por criar estruturas de dados denominadas índices, capazes de permitir que uma consulta seja realizada sem que seja necessário analisar toda a base de dados (MANNING; RAGHAVAN; SCHÜTZE, 2007). Os índices são semelhantes ao sumário de um livro, ou seja, uma lista detalhada, com a indicação de localização no texto e dos principais tópicos abordados no interior deste. Os índices foram utilizados para otimizar a velocidade e o desempenho da busca por um documento relevante em relação aos termos buscados. O custo pelo ganho de tempo durante a recuperação de informação é o espaço de armazenamento computacional adicional necessário para armazenar o índice.

Para analisar de forma descritiva os termos mais frequentes, foram utilizadas as frequências absoluta e relativa. Além disso, para analisar as notícias de forma geral, foram feitas nuvens de palavras (FELLOWS, 2012). As nuvens de palavras são um meio de representar a frequência que as palavras aparecem em um texto, de forma que quanto maior e quanto mais escura a palavra estiver na figura, maior é a frequência da mesma.

Para avaliar as palavras em relação aos anos de publicação das notícias, foi utilizada uma nuvem de comparação (FELLOWS, 2012). A nuvem de comparação contrasta os termos com maiores frequências em diferentes grupos, mostrando as diferentes opiniões/concepções por grupo. Ainda, foi realizada uma Análise de Correspondência (MINGOTI, 2005) para construção dos mapas perceptuais. A concepção geral da Análise de Correspondência (AC), entre outros aspectos, é que nela se permite a inclusão de variáveis categóricas, apropriadas para dados nominais. Recebe o nome de Análise de Correspondência pelo fato de estar interessada em estudar a correspondência entre as variáveis.

Conforme Carvalho e Struchiner (1992), este método permite a visualização gráfica das relações mais importantes de um grande conjunto de variáveis entre si (categóricas e contínuas categorizadas). A AC parte de uma matriz de dados representados por uma Tabela de Contingência e converge para um gráfico que exhibe as linhas e as colunas da matriz como pontos de um espaço vetorial de dimensão menor que a original, de maneira a estabelecer relações entre linhas, colunas, e entre linhas e colunas, que possam ser interpretáveis (GREENACRE; HASTIE, 1984). A partir dos princípios geométricos da AC, é possível representar dentro do Espaço

Euclidiano as distâncias entre os pontos linha e/ou coluna resultantes da associação entre as variáveis da tabela de contingência. Assim, tem-se o gráfico denominado mapa de correspondência ou mapa perceptual, que facilita a visualização das relações existentes entre as variáveis (LOURENÇO, 1997).

Para avaliar a percepção das notícias sobre o assunto principal, foi utilizada uma análise de sentimentos. Segundo Liu (2010), a análise de sentimentos, ou mineração de opinião, é o estudo computacional de opiniões, sentimentos e emoções expressas em texto. A análise de sentimentos essencialmente tenta inferir os sentimentos das pessoas com base em suas expressões de idioma. É uma das áreas de pesquisa mais ativas em processamento de linguagem natural e é amplamente estudada em mineração de dados, mineração de *Web* e mineração de texto. Na verdade, esta pesquisa ultrapassou a ciência da computação, chegando às ciências da gestão e às ciências sociais devido a sua importância para as empresas e para a sociedade como um todo. Segundo Liu (2012), a detecção do sentimento em um texto pode ocorrer em diferentes granularidades, sendo que a decisão do nível está sujeita ao contexto e à aplicação. Já a análise pode ser em nível de documento, sentença ou frases.

Com o intuito de agrupar textos e palavras similares, foi realizada uma Análise Hierárquica de Agrupamento (HAIR et al., 2009) via Método de Ward e distância Euclidiana. Os algoritmos hierárquicos são divididos em aglomerativos e divisivos. No método aglomerativo, considera-se, inicialmente, cada objeto distinto de todos os outros, ou seja, o número de *clusters* inicial é igual ao número de objetos. Em seguida, os objetos que tem menor distância, ou maior semelhança, vão sendo agrupados gradativamente, até ser formado apenas um grupo que contenha todos os objetos. O Método de Ward busca formar grupos de maneira a atingir sempre o menor erro interno entre os vetores que compõem cada grupo e o vetor médio do grupo, ou seja, o método busca o mínimo desvio-padrão entre os dados de cada grupo. Além disso, o Método de Ward utiliza como distância a soma dos quadrados entre os dois agrupamentos, feita sobre todas as variáveis. Este método forma grupos de maneira a minimizar a soma interna de quadrados, o que equivale a buscar o mínimo desvio-padrão entre os dados de cada grupo. A partir destas análises, foi possível fazer proposições relacionadas aos objetivos deste trabalho.

4 ANÁLISE DOS RESULTADOS

Um dos principais objetivos da mineração de textos é a quantificação e a caracterização de seu objeto de estudo. Com o levantamento dos números intrínsecos à base de dados textuais, pode-se compreender a sua abrangência e iniciar a construção de inferências que antes estavam ocultas na forma de texto. A análise dos resultados foi realizada à medida que o estudo foi evoluindo e, com base nos achados, foram criadas proposições que estão alinhadas aos objetivos deste trabalho.

4.1 APRESENTAÇÃO DOS RESULTADOS

4.1.1 Análise geral

Na análise geral, foi realizada a pesquisa dos 100 termos mais frequentes encontrados nas notícias extraídas. Além disso, foi criada uma nuvem de palavras mostrando de forma gráfica esta primeira análise. Com base nessa primeira análise, foi possível determinar uma série de achados interessantes.

Foram listadas as 100 palavras mais frequentes entre os 1.158.304 de termos encontrados após o pré-processamento nas 2.605 notícias analisadas. Logo, pode-se avaliar que as 5 palavras mais utilizadas nos textos foram: *blockchain*, *technology*, *bitcoin*, *bank* e *transaction*. Juntas, as 5 palavras representam 6,26% dos termos utilizados em todos os textos. Adicionalmente, observa-se que a palavra *blockchain* é largamente utilizada em todos os textos, sendo que foi utilizada 2,43 vezes mais que a segunda palavra mais frequente, *technology*. Ao todo, as 100 palavras mais frequentes, conforme mostrado na Tabela 1, representam 24,53% dos termos em todos os textos.

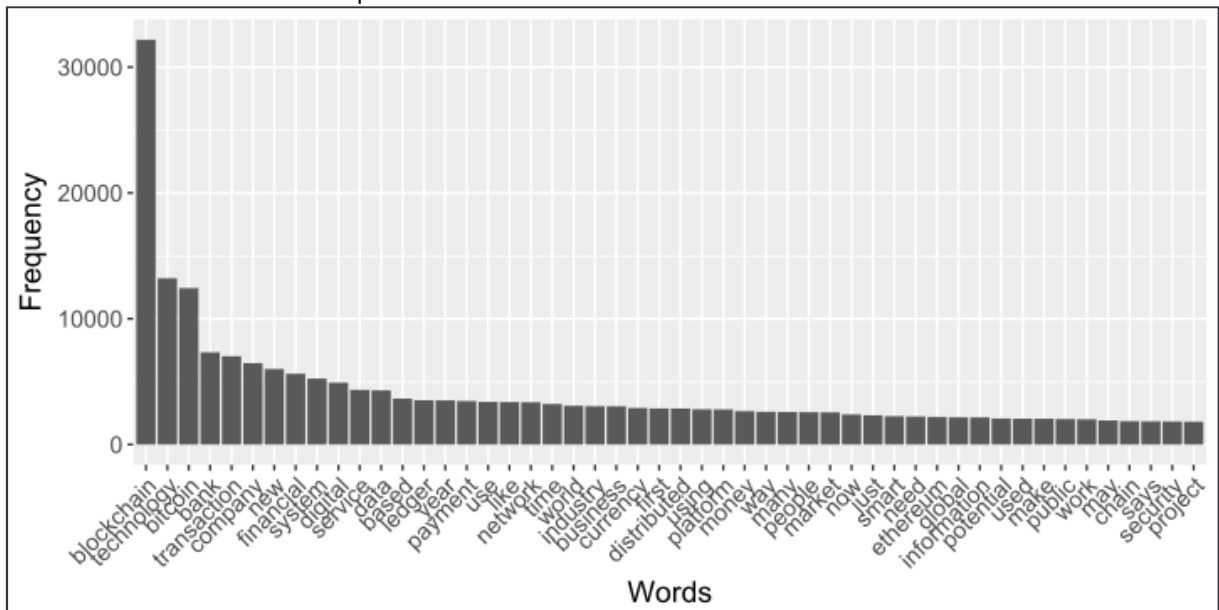
Tabela 1 – Palavras mais frequentes

Word	N	%	Word	N	%
<i>blockchain</i>	32184	2,79%	contracts	1805	0,16%
technology	13221	1,15%	exchange	1789	0,16%
bitcoin	12438	1,08%	cryptocurrency	1776	0,15%
bank	7332	0,64%	get	1753	0,15%
transaction	7025	0,61%	open	1739	0,15%
company	6474	0,56%	internet	1736	0,15%
new	6009	0,52%	value	1709	0,15%
financial	5632	0,49%	applications	1705	0,15%
system	5250	0,46%	trade	1675	0,15%
digital	4925	0,43%	well	1667	0,14%
service	4348	0,38%	see	1658	0,14%
data	4309	0,37%	IBM	1657	0,14%
based	3653	0,32%	real	1656	0,14%
ledger	3524	0,31%	private	1650	0,14%
year	3516	0,30%	future	1644	0,14%
payment	3461	0,30%	even	1639	0,14%
use	3389	0,29%	group	1611	0,14%
like	3378	0,29%	government	1550	0,13%
network	3360	0,29%	around	1522	0,13%
time	3223	0,28%	process	1505	0,13%
world	3088	0,27%	banking	1503	0,13%
industry	3050	0,26%	users	1497	0,13%
business	3043	0,26%	without	1495	0,13%
currency	2908	0,25%	secure	1463	0,13%
first	2868	0,25%	trading	1456	0,13%
distributed	2864	0,25%	according	1449	0,13%
using	2801	0,24%	two	1447	0,13%
platform	2790	0,24%	million	1439	0,12%
money	2655	0,23%	help	1437	0,12%
way	2601	0,23%	including	1436	0,12%
many	2588	0,22%	still	1432	0,12%
people	2574	0,22%	much	1427	0,12%
market	2560	0,22%	create	1423	0,12%
now	2395	0,21%	decentralized	1416	0,12%
just	2331	0,20%	record	1399	0,12%
smart	2257	0,20%	finance	1395	0,12%
need	2231	0,19%	working	1391	0,12%
ethereum	2195	0,19%	institutions	1366	0,12%
global	2167	0,19%	development	1364	0,12%
information	2163	0,19%	things	1325	0,11%
potential	2068	0,18%	central	1322	0,11%
used	2054	0,18%	every	1293	0,11%
make	2050	0,18%	take	1289	0,11%
public	2024	0,18%	tech	1273	0,11%
work	2009	0,17%	com	1261	0,11%
may	1916	0,17%	last	1257	0,11%
chain	1856	0,16%	trust	1251	0,11%
says	1845	0,16%	report	1246	0,11%
security	1833	0,16%	software	1240	0,11%
project	1816	0,16%	number	1235	0,11%

Fonte: Elaborado pelo Autor (2018).

Analisando a lista das 100 palavras mais frequentes, é possível observar a ocorrência de termos referentes a tecnologias de *blockchain* (*bitcoin*, *ethereum*); relacionados ao tipo de *blockchain* (*public*, *private*); bem como representantes de setores da indústria e instituições (*bank*, *financial*, *government*). Adicionalmente, também são encontrados termos referentes a casos de uso ou implementações da tecnologia, como *payment*, *currency*, *exchange*, *cryptocurrency* e *trade*. O Gráfico 1 mostra a frequência das 50 palavras mais utilizadas nos textos.

Gráfico 1 – Palavras mais frequentes



Fonte: Elaborado pelo Autor (2018).

A Figura 5 apresenta uma nuvem das palavras mais utilizadas nos textos. Cabe ressaltar que o tamanho das palavras é proporcional à frequência de utilização.

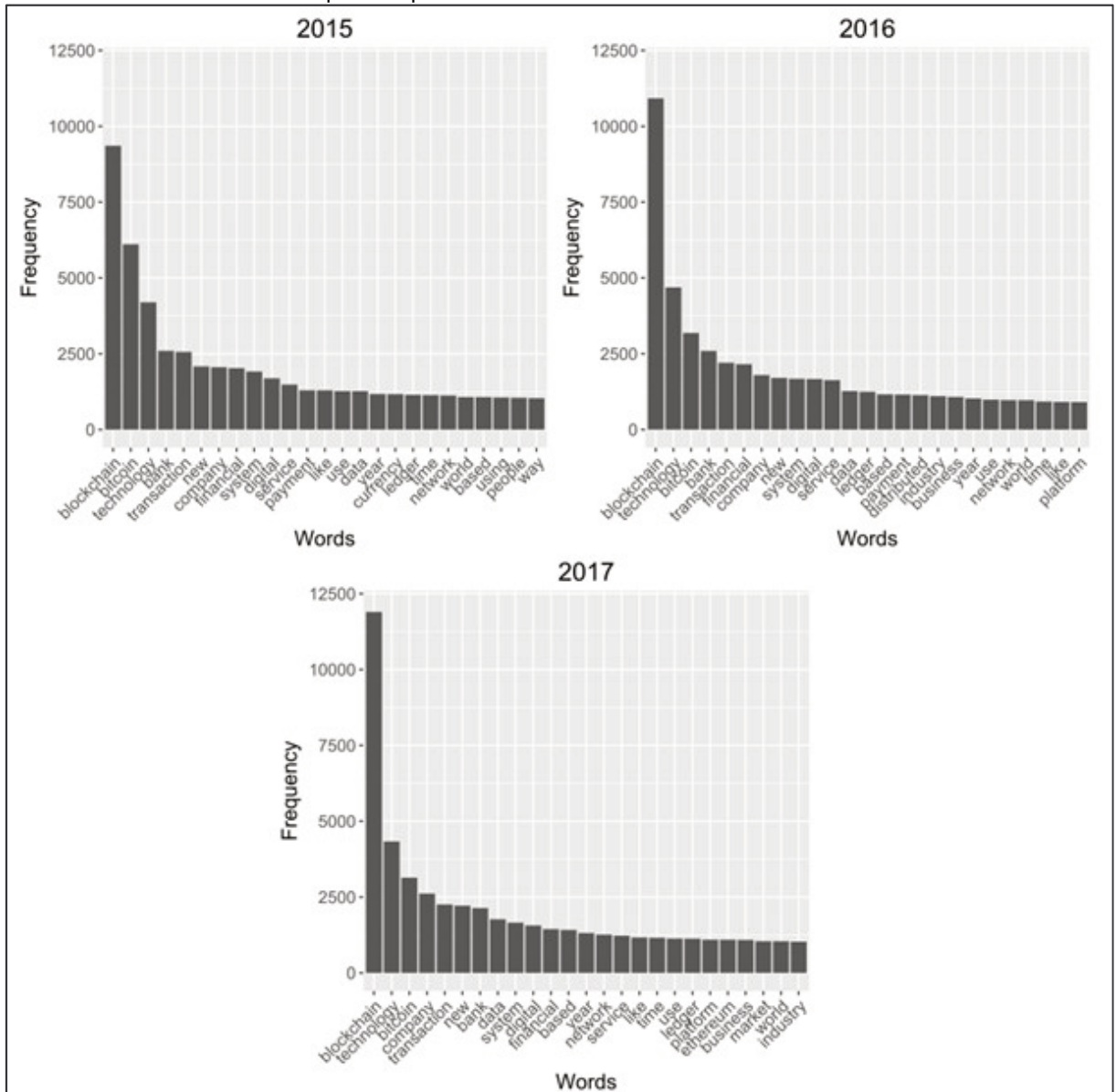
Tabela 2 – Palavras mais frequentes por ano

2015			2016			2017		
Word	N	%	Word	N	%	Word	N	%
blockchain	9359	2,33%	blockchain	10923	3,16%	blockchain	11902	2,93%
bitcoin	6111	1,52%	technology	4689	1,36%	technology	4334	1,07%
technology	4198	1,05%	bitcoin	3187	0,92%	bitcoin	3140	0,77%
bank	2598	0,65%	bank	2596	0,75%	company	2616	0,64%
transaction	2563	0,64%	transaction	2205	0,64%	transaction	2257	0,56%
new	2085	0,52%	financial	2158	0,62%	new	2218	0,55%
company	2059	0,51%	company	1799	0,52%	bank	2138	0,53%
financial	2028	0,50%	new	1706	0,49%	data	1769	0,44%
system	1916	0,48%	system	1677	0,49%	system	1657	0,41%
digital	1692	0,42%	digital	1668	0,48%	digital	1565	0,39%
service	1485	0,37%	service	1633	0,47%	financial	1446	0,36%
payment	1296	0,32%	data	1272	0,37%	based	1417	0,35%
like	1295	0,32%	ledger	1249	0,36%	year	1315	0,32%
use	1269	0,32%	based	1164	0,34%	network	1262	0,31%
data	1268	0,32%	payment	1156	0,33%	service	1230	0,30%
year	1174	0,29%	distributed	1137	0,33%	like	1169	0,29%
currency	1170	0,29%	industry	1103	0,32%	time	1160	0,29%
ledger	1146	0,29%	business	1077	0,31%	use	1133	0,28%
time	1137	0,28%	year	1027	0,30%	ledger	1129	0,28%
network	1125	0,28%	use	987	0,29%	platform	1096	0,27%
world	1073	0,27%	network	973	0,28%	ethereum	1095	0,27%
based	1072	0,27%	world	971	0,28%	business	1087	0,27%
using	1060	0,26%	time	926	0,27%	market	1045	0,26%
people	1051	0,26%	like	914	0,26%	world	1044	0,26%
way	1038	0,26%	platform	908	0,26%	industry	1029	0,25%

Fonte: Elaborado pelo Autor (2018).

Observa-se que as palavras mais frequentes foram muito similares entre os anos, podendo-se destacar as seguintes diferenças: dos 466 arquivos de notícias, 167 eram referentes ao ano de 2015, 144 eram de 2016 e os demais 155 eram de 2017. Em 2015, foram encontrados 401.594 termos; em 2016, foram 345.745; e, em 2017, foram encontrados 405.734 termos. As palavras *currency*, *payment* e *people* figuram entre as mais frequentes em 2015 e não aparecem entre as mais frequentes em 2016 e 2017. As palavras *business*, *industry* e *platform* figuram entre as mais frequentes nos anos de 2016 e 2017 e não aparecem entre as mais frequentes em 2015. A palavra *ethereum* figura entre as mais frequentes em 2017 e não aparece entre as mais frequentes em 2015 e 2016. O Gráfico 2 apresenta as palavras mais frequentes por ano.

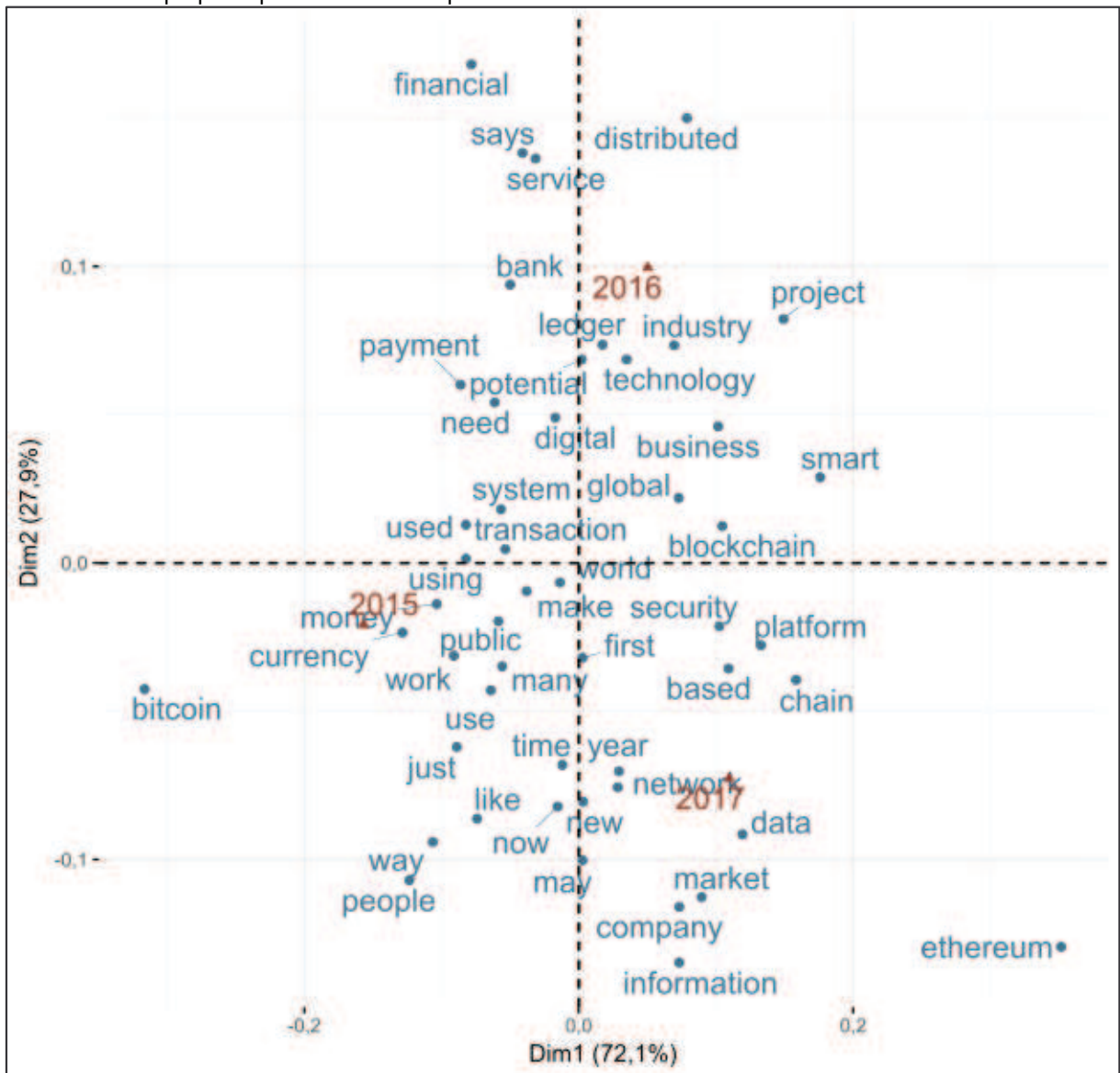
Gráfico 2 – Palavras mais frequentes por ano



Fonte: Elaborado pelo Autor (2018).

Logo após, foi feita uma nuvem de comparação. As nuvens de comparação são utilizadas para comparar as frequências que os termos aparecem entre grupos, contrastando palavras com maiores frequências em cada um. A Figura 6 apresenta uma nuvem de comparação entre os anos de 2015, 2016 e 2017.

Gráfico 3 – Mapa perceptual dos termos por ano



Fonte: Elaborado pelo Autor (2018).

Neste gráfico, pode-se observar que alguns dos termos mais associados ao ano de 2015 foram *money*, *currency*, *work*, *public*, *using* e *transaction*. Alguns dos termos mais associados ao ano de 2016 foram *ledger*, *industry*, *bank*, *technology*, *potential* e *service*. Alguns dos termos mais associados ao ano de 2017 foram *data*, *network*, *new*, *market*, *year* e *company*.

4.1.3 Análise de sentimentos

Segundo Liu (2010), a análise de sentimentos, ou mineração de opinião, é o estudo computacional de opiniões, sentimentos e emoções expressas em texto.

Ainda, ao discutir sentimentos subjetivos de emoção ou opiniões, é importante diferenciar duas noções: os estados mentais das pessoas e a linguagem utilizada para descrevê-los. Existem 6 tipos de emoções primárias (amor, alegria, surpresa, raiva, tristeza e medo). Entretanto, existem inúmeras expressões que podem descrever tais emoções. Da mesma forma, há também um grande número de expressões de opinião que descrevem sentimentos positivos ou negativos. A análise de sentimentos essencialmente tenta inferir os sentimentos das pessoas com base em suas expressões de idioma.

Dessa forma, palavras e expressões específicas foram pré-definidas como Sentimentos Positivos e Sentimentos Negativos por Liu (2010), por aparecerem dessa forma em estudos prévios. A análise de sentimento baseia-se no cruzamento das palavras encontradas nos textos com a matriz-referência de sentimentos.

Ao total, foram encontradas 82.653 correspondências entre as palavras nos textos e a matriz-referência, sendo que 54.605 foram correspondências referentes a Sentimentos Positivos e 28.048 foram correspondências referentes a Sentimentos Negativos. Em suma, o escore geral é de 26.557, o que indica um posicionamento positivo dos textos em relação ao tema principal. A Tabela 3 apresenta as 25 palavras mais frequentes dos Sentimentos Positivos e Negativos.

Tabela 3 – Palavras mais frequentes dos Sentimentos Positivos e Negativos

Sentimentos Positivos			Sentimentos Negativos		
	N	%		N	%
work	2603	4,77%	problem	1376	4,91%
well	1667	3,05%	issues	1165	4,15%
secure	1463	2,68%	risk	1151	4,10%
trust	1251	2,29%	fraud	496	1,77%
innovation	1168	2,14%	hard	491	1,75%
popular	971	1,78%	complex	410	1,46%
top	842	1,54%	difficult	359	1,28%
right	769	1,41%	limited	338	1,21%
support	766	1,40%	expensive	279	0,99%
important	765	1,40%	dangerous	268	0,96%
better	761	1,39%	critical	267	0,95%
benefits	694	1,27%	hype	267	0,95%
good	680	1,25%	vice	258	0,92%
best	644	1,18%	lack	255	0,91%
available	619	1,13%	impossible	246	0,88%
improve	579	1,06%	concerns	232	0,83%
leading	554	1,01%	disruption	216	0,77%
disrupt	542	0,99%	tamper	215	0,77%
significant	514	0,94%	sap	207	0,74%
free	499	0,91%	bad	180	0,64%
transparent	489	0,90%	slow	166	0,59%
great	481	0,88%	attack	155	0,55%
efficient	480	0,88%	hack	152	0,54%
trusted	453	0,83%	lost	152	0,54%
clear	431	0,79%	break	151	0,54%

Fonte: Elaborado pelo Autor (2018).

Logo, pode-se observar que as 5 palavras mais frequentes dos Sentimentos Positivos foram *work*, *well*, *secure*, *trust* e *innovation*. Juntas, as 5 palavras representam 14,93% das expressões que indicam Sentimentos Positivos nos textos. As 5 palavras mais frequentes dos Sentimentos Negativos foram *problem*, *issues*, *risk*, *fraud* e *hard*. Juntas, as 5 palavras representam 16,68% das expressões que indicam Sentimentos Negativos nos textos.

4.1.4 Análise de associação

Nesta análise, foi verificada a associação entre os 42 principais termos encontrados. Os valores estão entre 0 e 1, que é interpretado como a força de associação entre os termos. Por exemplo, se a correlação for igual a 1, significa que os dois termos foram usados em conjunto em 100% dos textos, enquanto que se a correlação for igual a 0,50, os dois termos foram usados em conjunto em 50% dos

textos. O limite mínimo de correlação foi adotado como 0,70. A Tabela 4 apresenta uma análise de associação entre os temas mais frequentes.

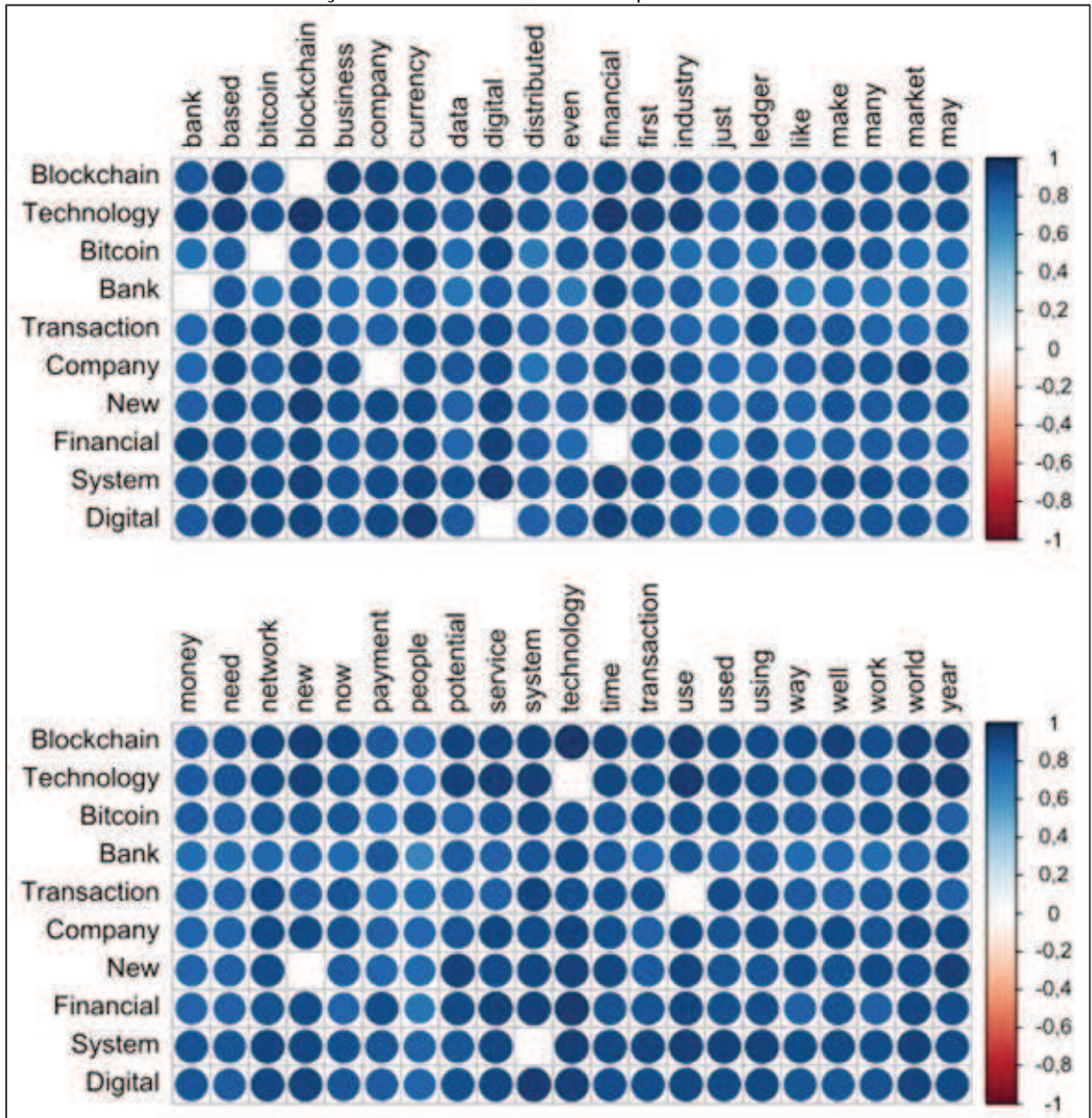
Tabela 4 – Análise de associação entre os termos mais frequentes

Words	<i>blockchain</i>	<i>technology</i>	<i>bitcoin</i>	<i>bank</i>	<i>transaction</i>	<i>company</i>	<i>new</i>	<i>financial</i>	<i>system</i>	<i>digital</i>
bank	0,84	0,89	0,75	-	0,79	0,78	0,81	0,90	0,85	0,83
based	0,94	0,93	0,83	0,84	0,88	0,90	0,89	0,88	0,91	0,91
bitcoin	0,84	0,87	-	0,75	0,87	0,83	0,86	0,86	0,89	0,90
blockchain	-	0,96	0,84	0,84	0,89	0,91	0,93	0,90	0,92	0,91
business	0,93	0,91	0,79	0,77	0,81	0,88	0,87	0,84	0,86	0,86
company	0,91	0,91	0,83	0,78	0,82	-	0,89	0,86	0,88	0,89
currency	0,88	0,90	0,91	0,84	0,87	0,86	0,89	0,89	0,91	0,94
data	0,87	0,83	0,76	0,72	0,85	0,84	0,80	0,79	0,87	0,83
digital	0,91	0,93	0,90	0,83	0,88	0,89	0,91	0,92	0,94	-
distributed	0,86	0,86	0,70	0,81	0,82	0,72	0,81	0,83	0,84	0,81
even	0,87	0,82	0,85	0,70	0,82	0,82	0,81	0,77	0,86	0,83
financial	0,90	0,95	0,86	0,90	0,86	0,86	0,88	-	0,92	0,92
first	0,93	0,93	0,88	0,83	0,85	0,90	0,91	0,87	0,90	0,89
industry	0,91	0,93	0,76	0,83	0,80	0,85	0,88	0,89	0,86	0,85
just	0,85	0,82	0,80	0,73	0,77	0,79	0,79	0,74	0,81	0,77
ledger	0,88	0,89	0,75	0,86	0,87	0,79	0,83	0,86	0,88	0,85
like	0,86	0,83	0,85	0,71	0,83	0,83	0,80	0,78	0,86	0,82
make	0,89	0,89	0,87	0,77	0,84	0,86	0,85	0,83	0,90	0,86
many	0,87	0,87	0,84	0,74	0,80	0,86	0,83	0,83	0,88	0,85
market	0,88	0,87	0,77	0,77	0,78	0,91	0,85	0,83	0,85	0,85
may	0,89	0,87	0,78	0,75	0,83	0,86	0,86	0,82	0,85	0,83
money	0,83	0,83	0,83	0,75	0,81	0,80	0,80	0,80	0,86	0,85
need	0,86	0,85	0,82	0,76	0,82	0,80	0,81	0,81	0,85	0,83
network	0,90	0,89	0,85	0,78	0,89	0,89	0,88	0,86	0,91	0,90
new	0,93	0,92	0,86	0,81	0,83	0,89	-	0,88	0,90	0,91
now	0,90	0,86	0,85	0,77	0,85	0,86	0,83	0,80	0,87	0,85
payment	0,83	0,86	0,77	0,84	0,78	0,81	0,80	0,88	0,85	0,83
people	0,81	0,79	0,85	0,65	0,77	0,79	0,77	0,72	0,81	0,80
potential	0,91	0,92	0,80	0,83	0,81	0,85	0,92	0,89	0,85	0,87
service	0,91	0,93	0,85	0,81	0,82	0,90	0,88	0,92	0,90	0,90
system	0,92	0,93	0,89	0,85	0,91	0,88	0,90	0,92	-	0,94
technology	0,96	-	0,87	0,89	0,87	0,91	0,92	0,95	0,93	0,93
time	0,92	0,90	0,85	0,84	0,87	0,87	0,90	0,86	0,90	0,88
transaction	0,89	0,87	0,87	0,79	0,87	0,82	0,83	0,86	0,91	0,88
use	0,94	0,95	0,88	0,85	-	0,90	0,91	0,91	0,93	0,90
used	0,90	0,90	0,87	0,81	0,88	0,86	0,86	0,87	0,92	0,89
using	0,88	0,89	0,86	0,83	0,88	0,88	0,85	0,88	0,92	0,90
way	0,89	0,86	0,84	0,76	0,83	0,87	0,88	0,81	0,89	0,86
well	0,92	0,90	0,84	0,79	0,81	0,89	0,86	0,86	0,89	0,85
work	0,87	0,85	0,87	0,75	0,84	0,85	0,89	0,82	0,88	0,86
world	0,93	0,93	0,89	0,81	0,87	0,90	0,88	0,90	0,92	0,91
year	0,94	0,93	0,82	0,87	0,81	0,90	0,93	0,88	0,89	0,89

Fonte: Elaborado pelo Autor (2018).

Neste estudo, todos os termos apresentaram alta associação entre si. Logo, na interpretação dos valores, correlações maiores que 0,90 foram consideradas altas. Dessa forma, pode-se destacar que a palavra *blockchain* apresentou alta associação com *based, business, company, digital, first, industry, new, potential, service, system, technology, time, use, well, world* e *year*. A palavra *technology* apresentou alta associação com *based, blockchain, business, company, digital, financial, first, industry, new, potential, service, system, use, world* e *year*. *Bitcoin* apresentou alta associação com *currency*. *Bank* não apresentou alta associação com nenhum dos termos, porém apresentou associação relevante com os termos *blockchain, technology, financial, system* e *digital*. *Transaction* apresentou alta associação com *system*. *Company* apresentou alta associação com *blockchain, market* e *technology*. *New* apresentou alta associação com *blockchain, digital, first, potential, technology, use* e *year*. *Financial* apresentou alta associação com *digital, service, system, technology* e *use*. *System* apresentou alta associação com *based, blockchain, currency, digital, financial, network, technology, transaction, use, used, using* e *world*. *Digital* apresentou alta associação com *based, blockchain, currency, financial, new, system, technology* e *world*. O Gráfico 4 ilustra os resultados.

Gráfico 4 – Análise de associação entre os termos mais frequentes



Fonte: Elaborado pelo Autor (2018).

Adicionalmente, foram feitas análises de associação com palavras selecionadas de acordo com os objetivos deste trabalho, de forma a verificar padrões e suportar a criação de proposições relacionadas. A Tabela 5 apresenta as 40 palavras com maior correlação com as palavras *strategy*, *use case*, *ethereum*, *hyperledger* e *industry*.

Tabela 5 – Análise de associação entre os termos específicos

strategy (n=300)	r	use case (n=716)	r	ethereum (n=2195)	r	hyperledger (n=622)	r	industry (n=3050)	r
focused	.83	poc	.87	build	.79	linux	.83	technology	.93
role	.82	<i>blockchain</i>	.86	ether	.78	fabric	.79	blockchain	.91
new	.81	project	.86	built	.76	ibm	.76	year	.90
part	.81	role	.86	two	.74	distributed	.74	service	.89
believes	.80	early	.85	move	.74	across	.72	financial	.89
move	.80	business	.84	called	.73	foundation	.72	new	.88
two	.79	part	.84	launched	.73	opening	.71	institutions	.88
digital	.79	distributed	.83	project	.73	maina	.71	provide	.88
early	.79	among	.83	designed	.73	ratification	.71	based	.87
potential	.79	potential	.83	option	.73	standardization	.70	business	.87
announced	.79	technology	.83	test	.73	tsc	.70	first	.87
built	.79	working	.83	users	.72	delays	.69	including	.87
startup	.79	benefits	.82	weeks	.72	framework	.68	number	.87
shift	.79	developed	.82	based	.71	project	.68	use	.87
firm	.79	development	.82	<i>blockchain</i>	.71	consortium	.68	among	.87
initiative	.79	applications	.82	developers	.71	ledger	.67	developing	.87
among	.78	existing	.82	development	.71	group	.67	existing	.87
last	.78	startup	.82	network	.71	regulatory	.67	world	.87
believe	.78	member	.81	among	.71	counterparties	.67	development	.86
lead	.78	new	.81	date	.71	nasty	.67	system	.86
blockchain	.77	process	.81	code	.71	member	.66	time	.86
building	.77	see	.81	launch	.71	standard	.66	end	.86
project	.77	service	.81	raised	.71	levet	.66	global	.86
today	.77	two	.81	autonomous	.71	members	.65	potential	.86
clients	.77	year	.81	organization	.71	manage	.65	working	.86
coinbase	.77	participants	.81	tied	.71	environment	.65	application	.86
designed	.77	international	.81	hard	.70	multi	.65	across	.85
funding	.77	based	.80	software	.70	test	.65	benefits	.85
innovation	.77	technical	.80	step	.70	effort	.65	currently	.85
offering	.77	time	.80	early	.70	reported	.65	distributed	.85
seek	.77	well	.80	given	.70	bnp	.65	however	.85
take	.77	area	.80	announced	.70	moving	.65	ledger	.85
usecase	.77	different	.80	decision	.70	developed	.64	many	.85
diretor	.77	enable	.80	moving	.70	development	.64	market	.85
previously	.77	followed	.80	building	.69	projects	.64	recently	.85
team	.77	given	.80	involved	.69	source	.64	solutions	.85
series	.77	industry	.80	projects	.69	maintenance	.64	used	.85
described	.77	partners	.80	tools	.69	vice	.64	well	.85
reveal	.77	processes	.80	year	.69	completed	.64	within	.85
business	.76	scale	.80	company	.69	pilot	.64	allow	.85

Fonte: Elaborado pelo Autor (2018).

Nesta análise, pode-se observar que as 5 palavras mais associadas com *strategy* foram *focused* ($r=0,83$), *role* ($r=0,82$), *new* ($r=0,81$), *part* ($r=0,81$) e *believes* ($r=0,80$). As 5 palavras mais associadas com o termo *use case* foram *Poc (proof of concept)* ($r=0,87$), *blockchain* ($r=0,86$), *project* ($r=0,86$), *role* ($r=0,86$) e *early* ($r=0,85$). As 5 palavras mais associadas com *ethereum* foram *build* ($r=0,79$), *ether* ($r=0,78$), *built* ($r=0,76$), *two* ($r=0,74$) e *move* ($r=0,74$). As 5 palavras mais associadas com

hyperledger foram *linux* ($r=0,83$), *fabric* ($r=0,79$), *IBM* ($r=0,76$), *distributed* ($r=0,74$) e *across* ($r=0,72$). As 5 palavras mais associadas com *industry* foram *technology* ($r=0,93$), *blockchain* ($r=0,91$), *year* ($r=0,90$), *service* ($r=0,89$) e *financial* ($r=0,89$).

4.1.5 Análise de agrupamento

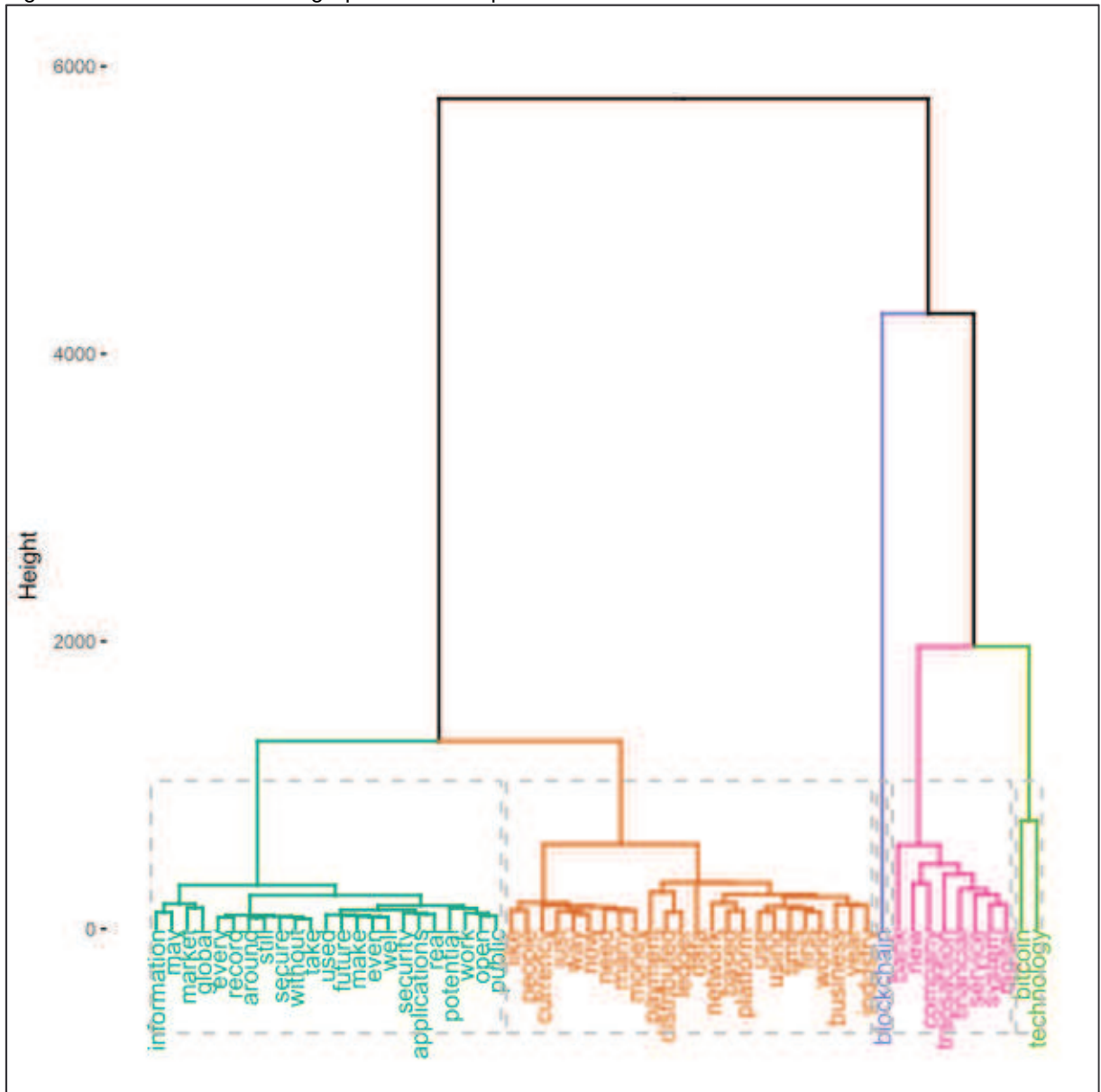
Com o intuito de agrupar textos e palavras similares, foi realizada uma Análise Hierárquica de Agrupamento (HAIR et al., 2009) via Método de Ward. O método de Ward busca formar grupos de maneira a atingir sempre o menor erro interno entre os vetores que compõem cada grupo e o vetor médio do grupo, ou seja, o método busca o mínimo desvio-padrão entre os dados de cada grupo.

A escolha do número de grupos foi feita a partir do dendograma. O dendograma é uma ferramenta apropriada para definir o número de grupos, pois uma boa classificação pode ser obtida ao se cortar o dendograma numa zona onde as separações entre classes correspondam a grandes distâncias (dissimilaridades).

4.1.5.1 Análise por palavra

Para agrupar as palavras por similaridade, foram utilizadas as 58 palavras mais frequentes nos textos. Foram consideradas as palavras que aparecem em, pelo menos, 60% de todos os textos. A Figura 7 apresenta o dendograma de agrupamento das palavras e se optou por trabalhar com cinco grupos.

Figura 7 – Dendrograma de agrupamento das palavras



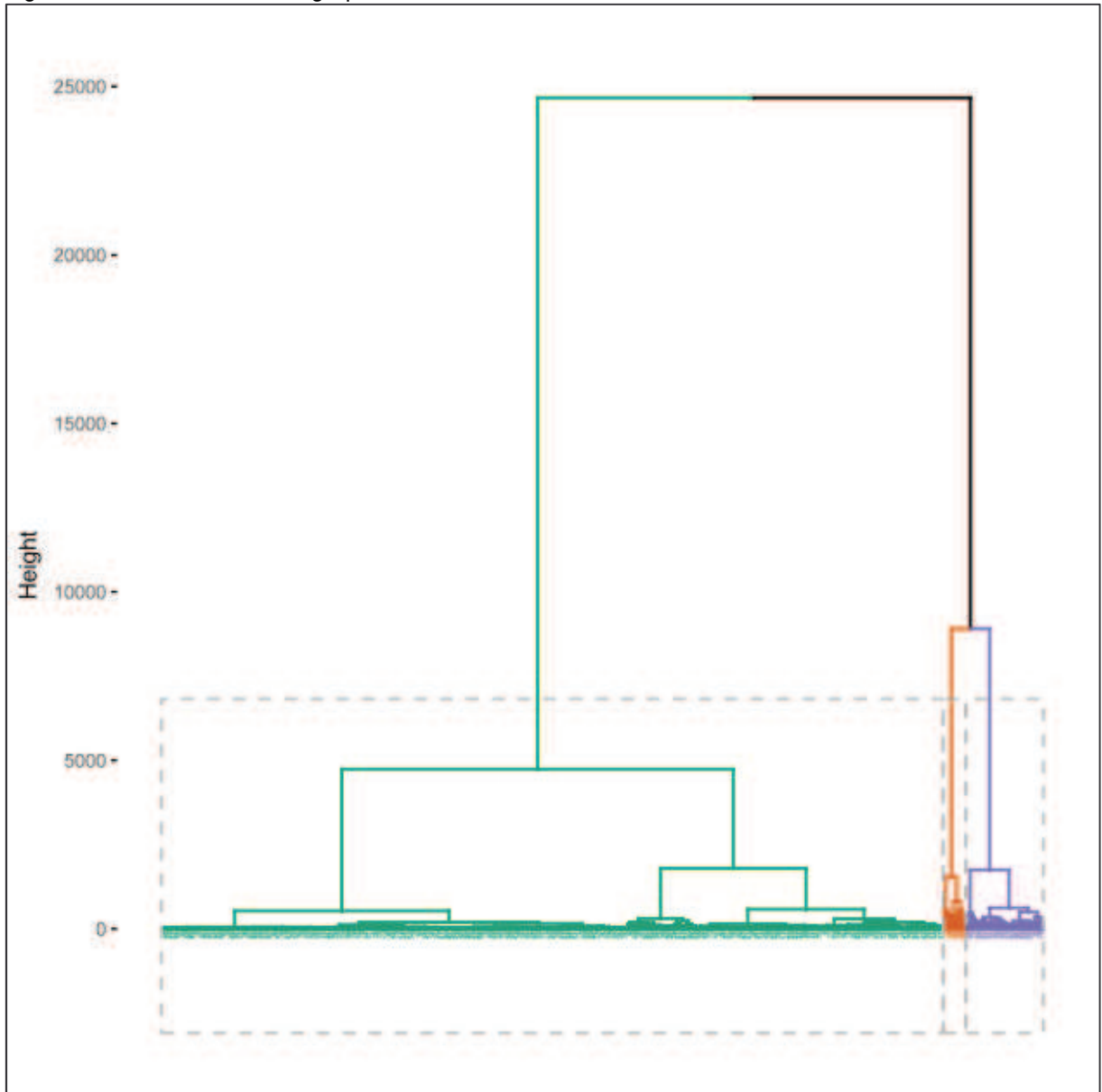
Fonte: Elaborado pelo Autor (2018).

Pode-se observar que o grupo em cor verde-água é composto pelas palavras *information, may, market, global, every, record, around, still, secure, without, take, used, future, make, even, well, security, applications, real, potential, open* e *public*. O grupo em cor laranja é composto pelas palavras *like, people, currency, just, way, now, need, money, payment, distributed, ledger, data, network, based, platform, use, using, time, first, world, business, year* e *industry*. A palavra mais frequente do estudo (*blockchain*) não é agrupada a nenhum outro termo devido a sua alta frequência entre os textos e está representada pela cor roxo. O grupo na cor rosa é composto pelas palavras *bank, new, company, transaction, financial, service, system* e *digital*. As palavras *bitcoin* e *technology* formam o último grupo, representado pela cor verde.

4.1.5.2 Análise por arquivos de texto

Nesta etapa, foi realizada uma análise de agrupamento dos 466 arquivos de texto que armazenam as notícias estudadas. A Figura 8 apresenta o dendograma de agrupamento dos textos e se optou por trabalhar com três grupos.

Figura 8 – Dendograma de agrupamento dos textos

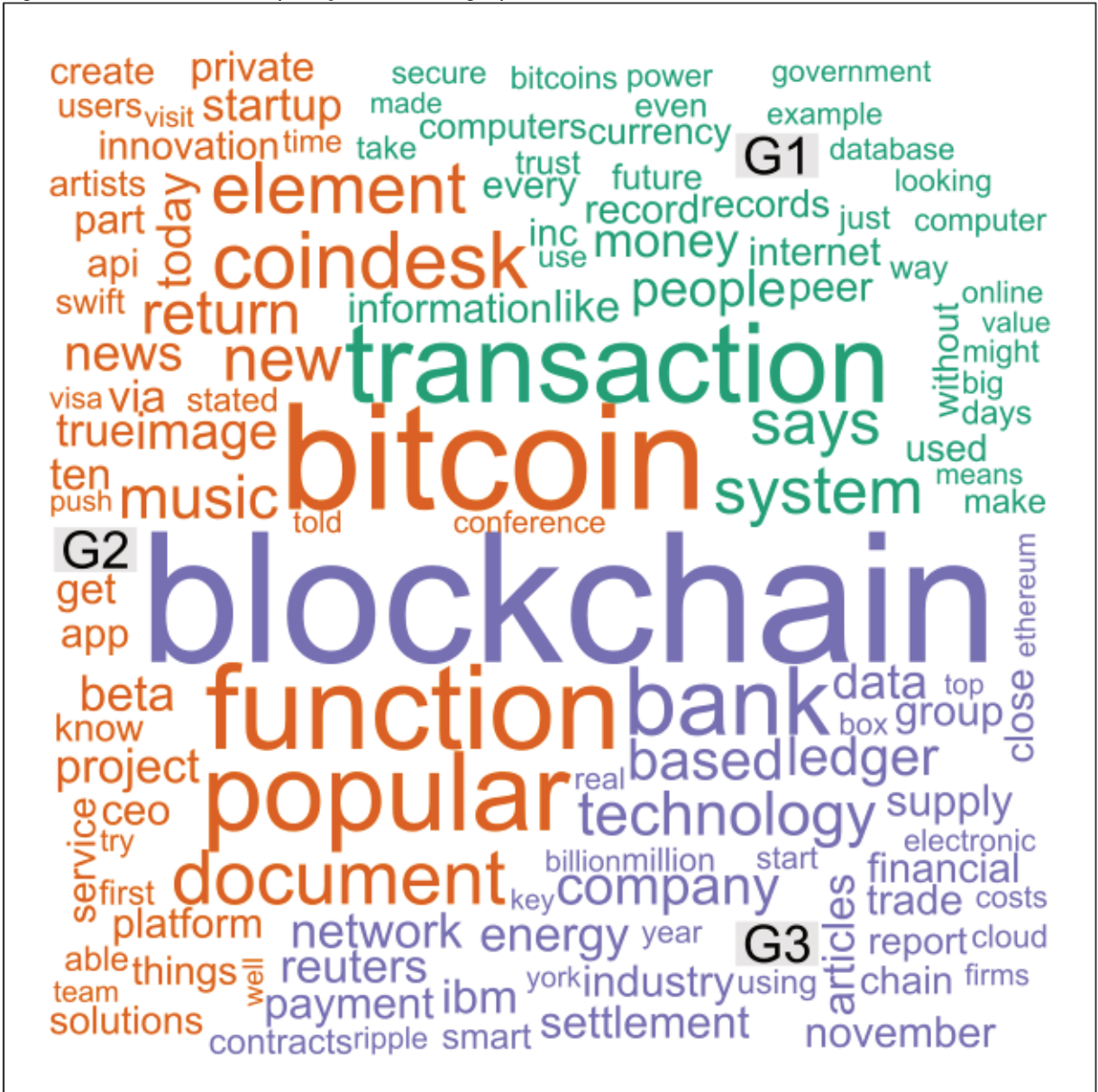


Fonte: Elaborado pelo Autor (2018).

Pode-se observar que grande parte dos textos foi agrupado em um único *cluster* devido à alta similaridade. O grupo representado na cor verde é composto por 413 arquivos (G1). O grupo na cor laranja engloba 12 arquivos (G2). O grupo na cor roxa possui 41 arquivos (G3). As nuvens de comparação são utilizadas para comparar

as frequências em que os termos aparecem entre grupos, contrastando palavras com maiores frequências em cada um. A Figura 9 apresenta uma nuvem de comparação entre os grupos.

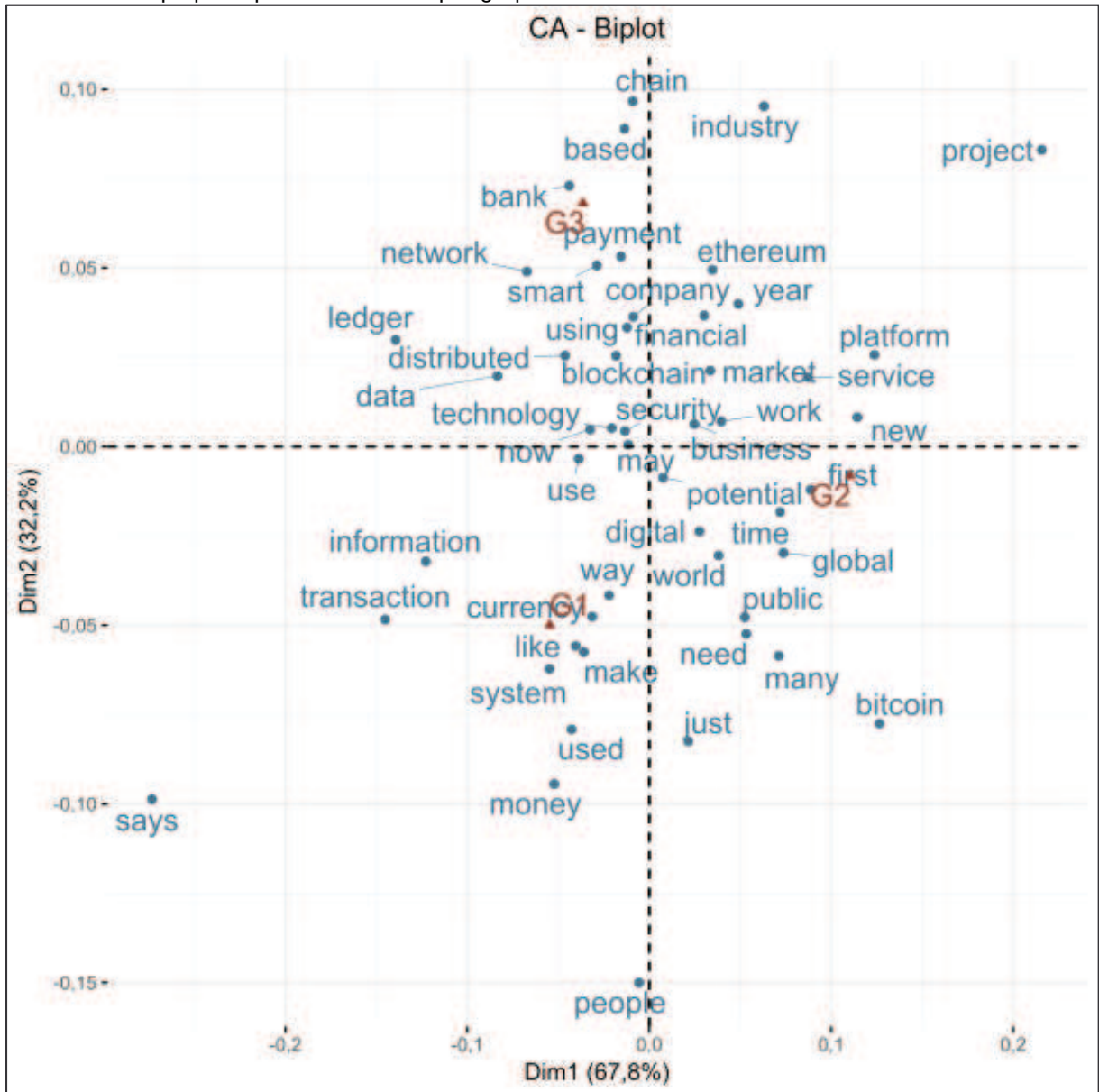
Figura 9 – Nuvem de comparação entre os grupos



Fonte: Elaborado pelo Autor (2018).

Pode-se observar que as palavras de maior destaque no Grupo 1 foram *transaction* e *system*. No Grupo 2, foram *bitcoin*, *function* e *popular*. Já no Grupo 3, foram *blockchain*, *bank* e *technology*. O Gráfico 5 apresenta o mapa percentual da relação entre os termos mais frequentes e o grupo. Quanto mais próximo o termo dos grupos, maior a associação.

Gráfico 5 – Mapa perceptual dos termos por grupo



Fonte: Elaborado pelo Autor (2018).

Logo, pode-se observar que alguns dos termos mais associados com o Grupo 1 foram *system*, *currency*, *make*, *like*, *information*, *used* e *money*. Alguns dos termos mais associados com o Grupo 2 foram *global*, *world*, *financial*, *public* e *need*. Alguns dos termos mais associados com o Grupo 3 foram *smart* (referente ao termo *smart contract*), *business*, *market*, *bank*, *blockchain* e *network*. Para testar a significância da relação, foi utilizado o teste Exato de Fisher (AGRESTI; KATERI, 2011). A Tabela 6 apresenta uma análise de associação entre os Grupos e os anos de publicação.

Tabela 6 – Análise de associação entre os Grupos e os anos de publicação

Ano / Grupo	G1		G2		G3		Valor-p ¹
	N	%	N	%	N	%	
2015	155	37,53%	5	41,67%	7	17,07%	0,088
2016	123	29,78%	4	33,33%	17	41,46%	
2017	135	32,69%	3	25,00%	17	41,46%	

Fonte: Elaborado pelo Autor (2018).

Nota: (1) Teste Exato de Fisher.

Verifica-se que houve associação marginalmente significativa (valor-p=0,088) entre ano e Grupo. A maioria dos textos do G1 e G2 foram publicados em 2015 (37,53% e 41,67%, respectivamente), enquanto que a maioria dos textos do G3 foram publicados em 2016 e 2017.

5 DISCUSSÃO DOS RESULTADOS

Blockchain é um banco de dados distribuído, compartilhado e criptografado que serve como um repositório público, irreversível e incorruptível de informações. Ele permite que pessoas não relacionadas possam chegar a um consenso sobre a ocorrência de uma determinada transação ou evento sem a necessidade de uma autoridade de controle. Segundo Drescher (2017), o fenômeno *blockchain* tem o potencial de mudar o modo como a sociedade opera.

De acordo com Wright e Filipp (2015), a tecnologia *blockchain* poderá apresentar a possibilidade de reduzir o papel de um dos atores econômicos e reguladores mais importantes em nossa sociedade, o intermediário. Segundo os autores, a *blockchain* tem o potencial de inaugurar uma nova era caracterizada por sistemas de pagamento globais, ativos digitais, governança descentralizada e, até mesmo, sistemas legais descentralizados. Isso permite que organizações coletivas e instituições sociais se tornem mais fluidas e promovam uma maior participação, transformando sobremaneira a forma como a governança corporativa e as instituições democráticas operam.

Na análise dos termos mais frequentes das notícias, pode-se verificar que os termos *decentralized* e *distributed* estão presentes dentre os mais frequentes, o que mostra que estas são características importantes desta tecnologia. Quando verificada a correlação do termo *distributed*, encontra-se forte associação com os termos *bank*, *financial* e *companies*. Com isso, identifica-se o interesse da indústria financeira e de instituições em relação aos benefícios que a descentralização que este fenômeno proporciona. Isso pode ser visto pela grande aparição dos seguintes termos nas notícias: *bank*, *financial*, *government* e *banking*.

O termo *bank* figura como o quarto termo mais mencionado em todas as notícias analisadas. Isso mostra que esse setor da indústria tem uma forte relação com o fenômeno deste estudo. Quando analisado o mapa percentual por ano referente ao ano de 2016, observa-se que o termo *bank* tem uma alta associação a este ano, o que mostra uma maior relação das notícias referentes aos bancos em relação ao fenômeno *blockchain* no ano de 2016. A palavra *bank* ainda apresentou correlação de 0,84 com o termo *blockchain*, o que reforça que este fenômeno está impactando diretamente o setor financeiro, mas especificamente os bancos. Além disso, quando verificada a correlação do termo *industry*, observa-se correlação de

0,89 com o termo *financeira*. Já o termo *financeira* apresenta forte correlação com *blockchain* e *technology*. Com isso, pode-se observar que este setor está focando os seus esforços estratégicos tecnológicos para estudar e criar soluções baseadas em *blockchain*. Com base nessa afirmação, surge a proposição P1:

Proposição P1: O fenômeno *blockchain* direciona as estratégias tecnológicas das empresas da indústria financeira para arquiteturas descentralizadas.

A teoria mostra que o fenômeno *blockchain* vem se transformando na medida em que o seu estágio de evolução progride. Esse aumento da maturidade se dá pelo crescimento do interesse das empresas nessa tecnologia.

Segundo Tapscot e Tapscot (2016), a *blockchain* que segue os princípios trazidos pela implementação do *bitcoin* é pública e de livre acesso. Além disso, os autores afirmam que o maior ganho com esta tecnologia está associado ao acesso aberto, removendo intermediários nos processos.

Entretanto, de acordo com Dinh et al. (2017), o interesse da indústria começou a impulsionar o desenvolvimento de novas plataformas *blockchain* projetadas para ambientes privados, nos quais os participantes são autenticados. Os sistemas *blockchain* em tais ambientes são chamados de privados (ou autorizados), ao contrário dos primeiros sistemas, que operam em ambientes públicos (ou sem permissão) onde qualquer um pode entrar e sair. Aplicações para negociação e liquidação de títulos, gestão de ativos e financeiros, operações bancárias e seguros estão sendo construídas e avaliadas utilizando essas tecnologias privadas. Atualmente, esses aplicativos são suportados por sistemas de banco de dados de nível corporativo, mas a *blockchain* tem o potencial de interromper esse *status quo* porque incorre em menores custos de infraestrutura e pessoas. Em particular, a imutabilidade e a transparência do *blockchain* ajudam a reduzir os erros humanos e a necessidade de intervenção manual devido a dados conflitantes. O *blockchain* pode ajudar a simplificar os processos de negócios, removendo esforços duplicados na governança de dados (DINH et al., 2017). Ainda segundo os autores, essas propriedades e os interesses comerciais das principais instituições bancárias e financeiras conferiram, às *blockchains* privadas, o potencial para aprimorar a prática atual no gerenciamento de dados.

Com base nos resultados das análises realizadas, foram encontradas nos textos referências para as seguintes tecnologias de *blockchain*: *bitcoin*, *ethereum*, *hyperledger*, *ripple* e *corda*. Nota-se que os termos *bitcoin* e *ethereum* figuram entre os 100 termos mais frequentes nas notícias. O termo *hyperledger* é o terceiro mais frequente nos textos referentes à tecnologia de *blockchain*, com 622 ocorrências, seguido pelo termo *ripple*, com 508 ocorrências e, por fim, o termo *corda*, com 156 ocorrências. É importante mencionar que as tecnologias *bitcoin* e *ethereum* são tecnologias de *blockchain* públicas. Já as tecnologias *hyperledger* e *corda* são *blockchains* privadas ou baseadas em permissões. Por fim a tecnologia *ripple* possui uma *blockchain* pública, porém baseada em permissões, e controlada por um consórcio de empresas.

Enquanto as tecnologias de tipo público são as que mais aparecem no conjunto total das notícias nos três anos analisados, nota-se um crescente aumento nas aparições das tecnologias privadas nas notícias dos anos de 2016 e 2017. A partir desta análise surge a proposição P2:

Proposição P2: O foco estratégico tecnológico das empresas impulsionou o desenvolvimento das tecnologias de *blockchain* privadas.

O pagamento entre países é um tópico do setor bancário que ainda obteve poucos benefícios do progresso recente na digitalização. A maioria das transações internacionais ainda são processadas usando um sistema bancário correspondente. No entanto, este sistema de entrada manual de transações não é uma solução escalável para o crescente espaço de pagamentos digitais. Embora isso possa parecer um problema para as instituições, um grande segmento da população pode se beneficiar dos avanços tecnológicos nesse setor.

De Meijer (2015) mostrou que o relatório do Tesouro do Reino Unido divulgado em março de 2015 é amplamente positivo em relação às criptografias digitais, observando a significativa promessa futura da tecnologia de *ledger* distribuída, que sustenta as criptomoedas digitais e seu potencial como tecnologia de pagamentos, ao mesmo tempo que destaca seu estado nascente e os riscos potenciais associados ao seu uso. Segundo o autor, a *blockchain* pode ser particularmente eficaz na melhoria dos pagamentos entre fronteiras, mais especificamente do correspondente bancário, dos pagamentos *business-to-business* e das remessas *peer-to-peer*. Em cada uma

dessas áreas, importantes ganhos de eficiência podem ser obtidos. A *blockchain* elimina a necessidade de bancos intermediários na esfera do banco correspondente, reduzindo assim as tarifas caras dos clientes, os atrasos no processamento de pagamentos e a falta de transparência. Na área de pagamentos *business-to-business*, a *blockchain* pode reduzir custos e tempo de transação, uma vez que levará o pagamento do remetente ao destinatário, evitando assim os numerosos intermediários de hoje. Também pode oferecer uma transferência de dinheiro de baixo custo para outro segmento crítico: o mercado de remessas *peer-to-peer*. Neste tipo de transferência, os trabalhadores migrantes internacionais enfrentam hoje um processo lento e caro para enviar dinheiro de volta para casa e para sua família.

De acordo com Robleh et al. (2014), a introdução de qualquer nova tecnologia permite repensar os processos de negócios associados à tecnologia anterior. No caso de pagamentos, quando os livros de papel foram primeiro informatizados, os processos subjacentes não foram alterados significativamente. Muitas vezes, a maior parte dos ganhos advindos da introdução de uma nova tecnologia não surge imediatamente, porque os processos que fazem uso da tecnologia também precisam ser repensados.

Segundo Guo e Liang (2016), os pagamentos interbancários muitas vezes dependem do processamento por empresas intermediárias de compensação, o que envolve uma série de processos complicados, incluindo escrituração contábil, reconciliação de transações, reconciliação de saldos, iniciação de pagamentos, entre outros. Portanto, o processo envolvido é demorado e caro. Usando pagamentos transacionais como exemplo, como os procedimentos de compensação para cada país são diferentes, uma remessa exige quase 3 dias para chegar ao destino. Isso demonstra a baixa eficiência e o imenso volume de recursos envolvidos.

Quando analisados os termos referentes aos sistemas de pagamentos entre países, observam-se os principais termos relacionados: *payment*, *exchange* e *trade*. É importante mencionar que os termos *exchange* e *trade* estão relacionados a casos de uso de sistemas de pagamentos entre países que foram amplamente mencionados nas notícias estudadas. Quando analisada a ocorrência das palavras por ano de coleta, pode-se observar que a palavra *payment* figura entre as 25 mais frequentes do ano de 2015, o que mostra que, neste ano, houve maior foco relacionado aos casos de uso baseados nos sistemas de pagamento. A palavra *payment* ainda aparece dentre as 25 mais listadas no ano de 2016. É importante mencionar que a palavra

payment apresenta alta correlação com as palavras *blockchain*, *technology*, *bank* e *company*, dentre outras. Com base nestas análises, surge a proposição P3:

Proposição P3: Sistemas de pagamento baseados em *blockchain* eliminam intermediários no processo de pagamentos entre países.

A tecnologia *blockchain* recebeu um aumento de interesse nos últimos anos, tanto da comunidade científica quanto da indústria. Hamida et al. (2017) afirma que essa tecnologia representa uma grande mudança de paradigma na forma como as soluções de cidades inteligentes serão construídas, operadas, consumidas e comercializadas em um futuro próximo. Mesmo que as *blockchains* tenham um tremendo impacto potencial nos negócios e nas sociedades, há muitos desafios abertos que precisam ser cuidadosamente enfrentados.

De acordo com Oh e Wallsten (2018), *blockchain* tem benefícios e custos e, como acontece com qualquer tecnologia, pode exacerbar ou mitigar as fraquezas das instituições existentes. A nova tecnologia tem implicações para a pesquisa em economia nova institucional e economia de custos de transação, particularmente em direitos de propriedade e estudos de corrupção. Uma vasta gama de atividades não relacionadas ao mercado pode ser medida, trocada e monitorada com *blockchain*. Novos economistas institucionais podem encontrar um novo mundo de economia de recursos para estudar. No entanto, a *blockchain* é melhor entendida em melhorias para a eficiência econômica e de governança.

Ainda segundo Oh e Wallsten (2018), o assunto é parcialmente revolucionário e parcialmente exagerado, levantando questões similares sobre a tecnologia e as instituições que foram levantadas antes do seu surgimento. Como tão poucas tecnologias revolucionárias acabam sendo assim, é improvável que a *blockchain* desmonte e substitua o estado moderno. No entanto, não se pode conhecer os efeitos por certo, e os maiores efeitos provavelmente serão de aplicações que não se pode prever. Entretanto, os pesquisadores enfrentam uma infinidade de oportunidades para ajudar a identificar e entender os efeitos econômicos imaginários e reais da *blockchain*.

Na teoria da difusão das inovações, Rogers (2003) define a difusão como o processo pelo qual (1) uma inovação (2) é comunicada através de certos canais (3)

ao longo do tempo (4) entre os membros de um sistema social. Os quatro principais elementos são a inovação, os canais de comunicação, o tempo e o sistema social.

Segundo Rogers (2003), uma inovação é uma ideia, prática ou objeto que é percebido como novo por um indivíduo ou outra unidade de adoção. Pouco importa, no que diz respeito ao comportamento humano, se uma ideia é ou não objetivamente nova, medida pelo lapso de tempo desde seu primeiro uso ou descoberta. A novidade percebida da ideia para o indivíduo determina sua reação a ela. Se uma ideia parece nova para o indivíduo, é uma inovação. A maioria das novas ideias cuja difusão foi analisada são inovações tecnológicas, e muitas vezes usa-se a palavra inovação e tecnologia como sinônimos. Uma tecnologia é um projeto de ação instrumental que reduz a incerteza nas relações de causa e efeito envolvidas na obtenção de um resultado desejado. Uma tecnologia geralmente tem dois componentes: (1) um aspecto de *hardware*, consistindo da ferramenta que incorpora a tecnologia como um objeto material ou físico, e (2) um aspecto de *software*, consistindo da base de informações para a ferramenta.

Ainda segundo Rogers (2003), a inovação deve ser amplamente adotada para se auto sustentar. Dentro da taxa de adoção, há um ponto em que uma inovação atinge massa crítica. As categorias de adotantes são: inovadores (*innovators*), adotantes iniciais (*early adopters*), maioria inicial (*early majority*), maioria tardia (*late majority*) e retardatários (*laggards*). A difusão manifesta-se de maneiras diferentes em várias culturas e campos e é altamente sujeita ao tipo de adotantes e ao processo de decisão da inovação.

A análise de sentimentos mostrou os principais termos positivos e negativos associados ao fenômeno estudado no conjunto de notícias. Dentre os termos positivos com maior incidência, foram selecionados termos relacionados a características importantes desta tecnologia, sendo eles: *secure, trust, innovation, benefits, available, transparent*. Na maioria das notícias, o fenômeno é descrito como um tipo de inovação disruptiva que oferece um ambiente seguro, confiável, disponível e transparente para a realização de transações de diversos tipos. Ainda, é possível concluir que a incidência de termos positivos é superior a de termos negativos sobre o tema, o que indica um posicionamento positivo dos textos frente ao fenômeno estudado.

Quando analisados os termos negativos encontrados nas notícias, percebe-se que ainda existem preocupações e incertezas relacionadas à *blockchain*. Os termos *problem* e *issues* são os termos negativos que mais aparecem nas notícias, o que

mostra que a tecnologia ainda está em desenvolvimento e que ainda gera incertezas. Analisando os textos, é possível encontrar diversas notícias tratando sobre o temor de fraudes associadas a esta tecnologia. Os termos *hard*, *complex* e *difficult* também aparecem dentre os principais termos encontrados e estão relacionados à complexidade para implementar soluções baseadas em *blockchain*, pois ainda existem poucos profissionais qualificados com esses conhecimentos. Também foi encontrado o termo *expensive*, que ressalta a necessidade de altos investimentos associados à implementação de soluções baseadas em *blockchain*. Ainda, se encontrou o termo *hype*, pois muitas notícias ainda tratam esse fenômeno como uma estratégia para enfatizar o impacto positivo desta tecnologia. Muitos ainda consideram que *blockchain* é um assunto que está na moda e que, por isso, é comentado por todo mundo.

Quando verificada a associação do termo *use case*, verifica-se correlação com os termos *poc* (*proof of concept*) e *early*. Isso mostra que a maioria dos casos de uso ainda estão relacionados a provas de conceitos e estão em estágio inicial de adoção. Com base nestas análises, surge a proposição P4:

Proposição P4: A tecnologia *blockchain* está no limiar das categorias *Innovators* e *Early Adopters* na perspectiva da difusão das inovações.

Um consórcio de bancos ganha em participação de mercado ao liquidar transações em tempo real (em vez de esperar três dias para que o negócio termine) e subscrever empréstimos em um dia (em vez de esperar duas semanas), tudo com risco mínimo. Segundo Plansky, O'Donnell e Richards (2016), os bancos também começam a executar operações cambiais no mesmo dia com taxas de câmbio ótimas, gastando uma fração dos custos exigidos no passado. Todas essas transações são rastreadas e as estatísticas são mantidas, para que os governos estejam cientes do movimento de capital através de suas fronteiras e para que a atividade seja monitorada por padrões que possam indicar lavagem de dinheiro. Contudo, a identidade dos comerciantes ou compradores individuais não é rastreável. Para Plansky, O'Donnell e Richards (2016), a tecnologia que poderia fazer tudo isso acontecer é a *blockchain*. Originalmente o nome formal do banco de dados de rastreamento subjacente ao *bitcoin*, o termo agora é usado amplamente para se referir

a qualquer livro eletrônico distribuído que usa algoritmos de *software* para registrar transações com confiabilidade e anonimato.

De fato, alguns bancos já estão explorando como o *blockchain* pode transformar suas abordagens de negociação e liquidação, operações de *backoffice* e investimento e gerenciamento de ativos de capital. Eles reconhecem que a tecnologia pode se tornar um fator diferenciador em suas próprias capacidades, permitindo que processem transações com mais eficiência, segurança, privacidade, confiabilidade e velocidade (PLANSKY; O'DONNELL; RICHARDS, 2016).

Segundo Wright e Fillipi (2015), a *blockchain* permite que pessoas não relacionadas cheguem a um consenso sobre a ocorrência de uma transação ou evento em particular sem a necessidade de uma autoridade controladora. Antes da invenção do *blockchain*, simplesmente não era possível coordenar atividades individuais pela Internet sem um corpo centralizado, garantindo que ninguém tivesse adulterado os dados. Um grupo de indivíduos não relacionados não podia confirmar que um evento ocorreu, sem depender de uma autoridade central para verificar se essa transação em particular não era fraudulenta ou inválida.

Conforme Kewell, Adams e Parry (2017), fundamentalmente, uma *blockchain* é um livro de transações de ativos digitais de quem possui o quê, quem transaciona o quê, do que é transacionado e quando é transacionado. As transações não são registradas em um único banco de dados, mas distribuídas nos computadores da rede de usuários (nós) do sistema. Nenhuma entidade única possui ou controla o *ledger* e, portanto, os membros da rede podem visualizar as transações registradas. As transações são gravadas e armazenadas em blocos, e cada bloco é vinculado cronologicamente e criptograficamente àqueles que o precedem para criar um registro imutável e resistente à falsificação. Todas as transações são marcadas com o tempo para fornecer um registro de quando as transações ocorreram e em que ordem, isso assegura contra o gasto duplo e adulteração de registros de transações anteriores. Quando um bloco é aceito pela rede e adicionado à cadeia, ele não pode ser alterado, passando a ser um registro permanente, transparente e imutável das transações.

O termo *transaction* é um dos termos mais frequentes nas notícias estudadas, totalizando 7.025 aparições. Quando analisados os termos por ano de publicação das notícias, nota-se que o termo *transaction* aparece com alta frequência em todos os anos estudados. Na análise do Mapa Percentual da relação entre os termos mais frequentes e o ano de publicação, nota-se que o termo está listado entre os termos

mais associados às notícias do ano de 2015. Na análise de associação, verifica-se que o termo tem correlação relevante com termos como *blockchain*, *business*, *company*, *financial*, *potential*, *industry* e *system*. Com base na discussão teórica sobre transações em *blockchain*, bem como pelos resultados das análises realizadas, surge a proposição P5:

Proposição P5: *Blockchain* tem potencial para afetar as transações tornando os dados acessíveis por meio de uma plataforma comum.

Uma variedade de desenvolvimentos em tecnologias de pagamento e moedas alternativas surgiram nos últimos anos. Algumas dessas inovações concentram-se em tornar os pagamentos mais acessíveis a uma ampla gama de usuários, como pagamentos por telefone celular, ao mesmo tempo em que dependem de uma entidade central confiável. De acordo com Robleh et al. (2014), inovações mais recentes introduziram uma estrutura descentralizada e fundamentalmente diferente nos sistemas de pagamento, baseando-se na criptografia em vez de uma autoridade central. Enquanto as moedas locais tecnicamente representam novas moedas, qualquer esquema desse tipo operando a uma taxa de câmbio fixa de um-para-um e respaldado por moedas nacionais mantém uma relação próxima com uma moeda existente.

Porém, é importante distinguir entre moedas digitais como sistemas de pagamento candidatos e moedas digitais como formas potenciais de dinheiro. Embora o *bitcoin* introduzisse uma nova moeda e uma nova tecnologia de pagamento juntas, a tecnologia de *blockchain* poderia, em teoria, ser usada sem a criação de uma nova moeda. Como enfatizado por Haldane e Qvigstad (2014), seria tecnicamente possível para um banco central existente emitir passivos apenas digitais em um sistema de pagamentos distribuídos de forma equivalente àqueles implantados pelas moedas digitais recentes.

Segundo Raskin e Yermack (2016), quando uma moeda digital autônoma circula em uma economia, ela compete com a moeda oficial emitida pelo banco central do país. A concorrência entre a moeda oficial e o dinheiro privado não é novidade e, em várias sociedades, o dinheiro alternativo incluiu mercadorias como o ouro e a prata, bem como outros bens que serviram como depósitos de valor e meios de troca.

De acordo com Shanbhogue (2011), em contraste com formas de dinheiro comumente usadas, como notas bancárias ou depósitos bancários, as moedas digitais não são uma reivindicação para ninguém. A este respeito, elas podem, portanto, ser consideradas como um tipo de mercadoria. Mas, ao contrário das *commodities* físicas, como o ouro, elas também são ativos intangíveis, ou *commodities* digitais. As moedas digitais só têm significado na medida em que os participantes concordam que têm significado. Esse acordo assume a forma de um livro-razão público e um processo de como as mudanças são feitas, incluindo a criação da nova moeda. Não ser uma responsabilidade do banco central (ou do Estado) não impede que moedas digitais sejam usadas como dinheiro, mas define uma diferença importante entre elas e as moedas nacionais.

Guo e Liang (2016) argumentam que o pagamento *peer-to-peer* também pode ser implementado usando a tecnologia *blockchain*, eliminando o elo intermediário de instituições financeiras terceiras, o que melhorará muito a eficiência do serviço e reduzirá os custos de transação dos bancos. Isso também permitirá que os bancos cumpram os requisitos de serviços de compensação de pagamentos rápidos e convenientes para atividades comerciais entre países. Segundo Guo e Liang (2016), a McKinsey fez uma estimativa que mostra que o custo de cada transação no negócio de pagamentos entre países pode ser bastante reduzido devido à aplicação de *blockchain* e moedas digitais.

Quando analisados os termos referentes a criptomoedas e sistemas de pagamentos entre países, observam-se os principais termos relacionados: *payment*, *currency*, *cryptocurrency*, *exchange* e *trade*. Na análise dos termos mais frequentes, os termos *currency* e *cryptocurrency* figuram dentre os mais frequentes nos textos. Na análise de associação, verifica-se que o termo *currency* possui alta correlação com termos como *blockchain*, *bitcoin* e *bank*. Com base nas proposições anteriores sobre transações e sistemas de pagamentos entre países, conectando esses achados à teoria acerca do tema de moedas digitais, conclui-se que as moedas digitais desempenham um papel importante na viabilização de transações para pagamentos entre países. Com base na discussão teórica e nas análises dos textos, surge a proposição P6:

Proposição P6: As moedas digitais aprimoram as transações em sistemas de pagamentos entre países.

As proposições geradas neste trabalho contribuem para a prática através dos achados relacionados às teorias estudadas. Por meio da pesquisa de descoberta de conhecimento nos textos, foi possível identificar padrões e achados que foram discutidos à luz das teorias da tecnologia *blockchain*, bem como da estratégia tecnológica. De forma prática, essas proposições podem ser usadas por empresas ou instituições que tem interesse em focar a sua estratégia tecnológica em *blockchain* pois, através desta pesquisa, foram identificados impactos do fenômeno *blockchain* na estratégia tecnológica.

Em relação à teoria, esta pesquisa possibilitou fazer uma análise discutindo dados empíricos, através das notícias, com as teorias existentes acerca do tema deste trabalho, gerando proposições como resultado das discussões teóricas. É importante mencionar que não foram encontrados, na literatura, trabalhos que buscassem trazer a análise aqui mostrada. As seis proposições que estão sintetizadas no Quadro 1, contribuindo para o conhecimento teórico e, de forma prática, para a estratégia tecnológica.

Quadro 1 – Proposições

Nº	Proposição
P1	O fenômeno <i>blockchain</i> direciona as estratégias tecnológicas das empresas da indústria financeira para arquiteturas descentralizadas.
P2	O foco estratégico tecnológico das empresas impulsionou o desenvolvimento das tecnologias de <i>blockchain</i> privadas.
P3	Sistemas de pagamento baseados em <i>blockchain</i> eliminam intermediários no processo de pagamentos entre países.
P4	A tecnologia <i>blockchain</i> está no limiar das categorias <i>Innovators</i> e <i>Early Adopters</i> da teoria da difusão das inovações.
P5	<i>Blockchain</i> tem potencial para afetar as transações tornando os dados acessíveis por meio de uma plataforma comum.
P6	As moedas digitais aprimoram as transações em sistemas de pagamentos entre países.

Fonte: Elaborado pelo Autor (2018).

6 CONSIDERAÇÕES FINAIS

O ambiente vem impondo às organizações novos desafios. No que tange aos aspectos tecnológicos, os produtos têm reduzido seus ciclos de vida continuamente, exigindo das empresas o reconhecimento e análise das novas oportunidades tecnológicas. Diante de tais desafios, a estratégia tecnológica surge como um auxílio no processo decisório relativo às tecnologias a serem adotadas, levando-se sempre em consideração as tendências tecnológicas e as necessidades do mercado.

Contratos, transações e os seus registros estão entre as estruturas definidoras dos sistemas econômicos, legais e políticos. Eles protegem ativos e estabelecem limites organizacionais. Eles determinam e verificam identidades e eventos crônicos. Eles governam as interações entre nações, organizações, comunidades e indivíduos. Eles guiam a ação gerencial e social, mas essas ferramentas críticas e as burocracias formadas para gerenciá-las não acompanharam a transformação digital da economia. Em um mundo digital, a maneira como regula-se e mantêm-se o controle administrativo necessita mudanças. A tecnologia de *blockchain* é referenciada por muitos como a possível solução para esta necessidade de mudança proporcionada pela economia digital.

O objetivo principal deste trabalho foi compreender de que forma o fenômeno *blockchain* tem afetado a estratégia tecnológica. Os objetivos específicos foram:

- a) Compreender quais os impactos foram gerados pelas implementações de *blockchain* na estratégia tecnológica nos diversos setores;
- b) Mapear como o foco estratégico tecnológico das empresas está impulsionando o desenvolvimento desta tecnologia;
- c) Identificar o estágio de maturidade desta tecnologia.

Com isso, foi realizada uma pesquisa exploratória com abordagem quantitativa, utilizando técnicas de Descoberta de Conhecimento em Textos, com o uso de ferramentas de mineração de textos. Primeiramente, foram realizados os procedimentos relacionados à coleta das notícias na *Web* e, posteriormente, foram realizados os procedimentos de preparação dos textos para então realizar a fase de análise através de técnicas de mineração de textos. Após a análise dos resultados, foi possível realizar uma discussão com a teoria, utilizando os resultados obtidos pelo processo de DCT e criando proposições como resultado deste estudo.

Em relação aos resultados obtidos neste trabalho, pôde-se verificar diversas contribuições gerenciais e acadêmicas. Dentre as contribuições gerenciais, identificou-se que há impacto do fenômeno *blockchain* para a estratégia tecnológica, principalmente para as instituições financeiras, onde os bancos são os principais impactados por esta tecnologia. Foi possível verificar que a característica de arquitetura distribuída trazida pela tecnologia de *blockchain* despertou interesse nas empresas deste segmento. Além disso, verificou-se que o interesse das empresas em explorar a utilização desta tecnologia corroborou para o desenvolvimento das tecnologias de *blockchain* privadas, mostrando maior aderência à estratégia tecnológica das empresas. Foram encontradas evidências de que esta tecnologia traz benefícios quando utilizada na implementação de sistemas de pagamento entre países, diminuindo os intermediários, bem como reduzindo custos e melhorando processos relacionados. Quanto ao grau de maturidade, foi possível verificar que a tecnologia ainda está em seus estágios iniciais, quando analisada à luz da teoria da difusão das inovações de Rogers (2003). Em relação ao potencial da tecnologia de *blockchain*, identificou-se que está no registro de transações através desta tecnologia. Com isso, conclui-se que esta tecnologia pode afetar as transações da mesma maneira que o sistema de posicionamento global (*Global Positioning System*) mudou o transporte pessoal, tornando os dados acessíveis por meio de uma plataforma comum. Por fim, a pesquisa trouxe dados que mostraram que as moedas digitais têm um papel importante para a viabilização dos sistemas de pagamentos entre países.

Quanto às contribuições acadêmicas, foi possível realizar um confronto entre o que está sendo publicado nas notícias da *Web* sobre o fenômeno e as teorias acadêmicas da estratégia tecnológica e de *blockchain*. Com isso, foi possível realizar um mapeamento de como o fenômeno *blockchain* está impactando a estratégia tecnológica. Ainda, foi possível verificar como esse fenômeno vem sendo explorado no campo teórico. Nota-se que a maioria das publicações acadêmicas ainda é relacionada ao assunto das moedas digitais. Porém, identificou-se, durante a pesquisa, que estão surgindo artigos e trabalhos relacionados a outros usos desta tecnologia, o que mostra que a academia está dando ênfase para o assunto.

Por fim, foram criadas proposições alinhadas com a teoria acerca do tema que poderá ajudar outros pesquisadores na identificação de novas áreas de pesquisa, auxiliando em estudos futuros. Através desta pesquisa, foi possível identificar diversos

conhecimentos relacionados ao fenômeno *blockchain* e o seu impacto na estratégia tecnológica.

6.1 LIMITAÇÕES DA PESQUISA

Em relação aos fatores limitantes, ressalta-se que este estudo dependeu das técnicas de mineração de textos disponibilizadas pelo *software* utilizado e, portanto, os resultados são dependentes de eventuais problemas técnicos na sua construção ou devido à utilização de técnicas não usuais. Vale salientar que a falta de conhecimento e experiência do autor em ferramentas deste tipo fez com que alguns resultados não tivessem êxito. Nota-se que o processo de DCT utilizando ferramentas de mineração de textos oferece uma gama interessante de recursos para a descoberta de conhecimento em dados não-estruturados. Porém, ainda possui algumas limitações como, por exemplo, não possibilitar que termos compostos por mais de uma palavra sejam analisados nos termos mais frequentes.

Em relação à amostra de notícias extraídas na etapa de coleta, existiram limitações quanto à quantidade de notícias listadas por consulta, devido a limitações que o *software* de busca oferece, bem como a limitação de tempo hábil para a efetiva coleta e tratamento dos textos. É importante salientar que mesmo tendo realizado a extração dos textos das notícias através do uso de um *software* para esta finalidade, foram encontradas limitações e restrições nessa etapa. As limitações foram relacionadas às diferentes características de estrutura de dados dos sites onde as notícias foram extraídas, bem como pela quantidade de códigos de programação que eram extraídos em conjunto com os textos das notícias.

Outro fator que pode ter limitado os resultados é que os textos usados para descoberta de conhecimento não foram selecionados por critérios de qualidade, mas somente por período. Isto poderia introduzir algum tipo de ruído no processo. Acredita-se que, com uma seleção melhor destes textos, pode-se encontrar conhecimentos mais consistentes e, assim, melhorar os resultados. Além disto, acredita-se ser possível melhorar o processo utilizando um número maior de notícias. Um trabalho futuro poderá avaliar esta hipótese.

Devido a características da ferramenta de extração de notícias, optou-se por extrair as notícias publicadas por um mesmo portal de forma unificada e, desta forma, essas notícias foram armazenadas em arquivos únicos separados por portal de

notícias. Por exemplo, todas as notícias do ano de 2015 oriundas do portal *Reuters* foram armazenadas em um mesmo arquivo. Com isso, as 2.605 notícias foram distribuídas em 466 arquivos, ficando misturadas. Isso fez com que a análise de agrupamento por arquivos de texto não trouxesse resultados efetivos, pois o cenário ideal para esta análise seria o de ter cada notícia armazenada em um arquivo separado, totalizando então 2.605 arquivos. Como as notícias ficaram misturadas em arquivos centralizadores, os assuntos se misturaram e prejudicaram a análise e agrupamento de arquivos.

6.2 SUGESTÕES PARA TRABALHOS FUTUROS

Os resultados encontrados neste trabalho não encerram a totalidade de estudos relacionados ao impacto do fenômeno *blockchain* na estratégia tecnológica. Ao contrário, fomentam o desenvolvimento de novos estudos sobre o assunto. Isso ocorre porque esse fenômeno tecnológico ainda está em fase inicial e ainda não chegou na sua maturidade.

Como sugestões para estudos futuros, destaca-se a validação das proposições aqui criadas, de forma a verificar a confirmação ou não das mesmas. Outra oportunidade de estudo está ligada a uma análise focada no impacto deste fenômeno em relação às instituições governamentais, com foco nos bancos centrais.

Ainda, é possível a realização de estudos focados em outros casos de uso e modelos de negócio que esta tecnologia vem proporcionando, como os casos relacionados a *supply chain*, que vem ganhando maior destaque nos últimos tempos.

Neste trabalho, foram encontradas evidências da relevância das moedas digitais associadas ao fenômeno estudado. Portanto, uma oportunidade de pesquisa futura seria uma análise das regulações existentes acerca destas moedas.

Existe outra oportunidade na proposição de soluções para superar desafios e limitações da tecnologia *blockchain*, como problemas de performance ou escalabilidade existentes atualmente. Por fim, sugere-se avaliar a eficácia das soluções propostas com um critério de avaliação objetivo. Embora várias soluções para desafios e limitações tenham sido apresentadas, muitas delas são apenas sugestões breves e carecem de avaliações concretas sobre sua eficácia.

REFERÊNCIAS

- ACHESON, N. What is smart property, and what can it be used for? **Fintechblue**, 9 dez. 2015.
- AGRESTI, A.; KATERI, M. Categorical data analysis. In: LOVRIC, M. (Ed.). . **International Encyclopedia of Statistical Science**. Berlin: Springer Berlin Heidelberg, 2011. p. 206–208.
- AINSWORTH, C. Extract URLs from Google’s Web SERPs. [**Personal Blog**], 9 jun. 2015.
- ARANHA, C. N. **Uma abordagem de pré-processamento automático para mineração de textos em português: Sob o enfoque da inteligência computacional**. Tese (Doutorado em Engenharia Elétrica)—Rio de Janeiro: Pontifícia Universidade Católica do Rio de Janeiro, 2007.
- ARANHA, C.; PASSOS, E. A tecnologia de mineração de textos. **RESI - Revista Eletrônica de Sistemas de Informação**, v. 2, p. 1–8, 2006.
- ARASTI, M.; KHALEGHI, M.; NOORI, J. Corporate-level technology strategy and its linkage with corporate strategy in multi-business companies: IKCO case study. **Technological Forecasting and Social Change**, v. 122, p. 243–252, set. 2017.
- AU, Y. A.; KAUFFMAN, R. J. The economics of mobile payments: Understanding stakeholder issues for an emerging financial technology application. **Electronic Commerce Research and Applications**, v. 7, n. 2, p. 141–164, jun. 2008.
- AVLONITIS, G. J.; PAPASTATHOPOULOU, P. G.; GOUNARIS, S. P. An empirically-based typology of product innovativeness for new financial services: Success and failure scenarios. **Journal of Product Innovation Management**, v. 18, n. 5, p. 324–342, set. 2001.
- BARBER, S. et al. Bitter to better: How to make bitcoin a better currency. International Conference on Financial Cryptography and Data Security. **Anais...Berlim: 2012**
- BASTOS, V. M. **Ambiente de desboberta de conhecimento na Web para a língua portuguesa**. Tese (Doutorado em Engenharia)—Rio de Janeiro: Universidade Federal do Rio de Janeiro, 2006.
- BECK, R. et al. Blockchain: The gateway to trust-free cryptographic transactions. **Twenty-Fourth European Conference on Information Systems (ECIS)**, v. 5, n. 15, p. 1–14, 2016.
- BISHOP, K. et al. **Succeeding through service innovation: A service perspective for education, research, business and government**. [s.l: s.n.].
- BROCKHOFF, K.; PEARSON, A. Technical and marketing aggressiveness and the effectiveness of research and development. **IEEE Transactions on Engineering Management**, v. 39, n. 4, p. 318–324, 1992.

CARRILHO JÚNIOR, J. R. **Desenvolvimento de uma metodologia para mineração de textos**. Dissertação (Mestrado em Engenharia Elétrica)—Rio de Janeiro: PUC-Rio, 2007.

CARVALHO, M. S.; STRUNCHINER, C. J. Análise de correspondência: Uma aplicação do método à avaliação de serviços de vacinação. **Caderno de Saúde Pública**, v. 8, n. 3, p. 287–301, 1992.

CHERUBIN, P. F. A integração entre a estratégia de negócio e a estratégia tecnológica em software-houses: O caso da Educom e da Gescom. **EnANPAD**, p. 1–14, 2000.

CLARKE, K. et al. Technology strategy in UK firms. **Technology Analysis & Strategic Management**, v. 7, n. 2, p. 169–191, jan. 1995.

DE MEIJER, C. R. W. The UK and Blockchain technology: A balanced approach. **Journal of Payments Strategy & Systems**, v. 9, n. 4, p. 220–229, 2015.

DEEDS, D. L. The role of intensity, technical development and absorptive capacity in creating entrepreneurial wealth in high technology start-ups. **Journal of Engineering and Technology Management**, v. 18, n. 1, p. 29–47, mar. 2001.

DINH, T. T. A. et al. Blockbench. **Proceedings of the 2017 ACM International Conference on Management of Data - SIGMOD '17**, p. 1085–1100, 2017.

DRESCHER, D. **Blockchain basics: A non-technical introduction in 25 steps**. Frankfurt: Apress, 2017.

DUPONT, Q.; MAURER, B. Ledgers and Law in the Blockchain. **Kings Review**, p. 1–15, 2015.

EBECKEN, N. F.; LOPES, M. C. S.; COSTA, M. C. A. Mineração de textos. In: REZENDE, S. DE O. (Ed.). **Sistemas Inteligente**. Barueri: Manole, 2003. p. 337–370.

ECONOMIDES, N.; KATSAMAKAS, E. Two-sided competition of proprietary vs. open source technology platforms and the implications for the software industry. **Management Science**, v. 52, n. 7, p. 1057–1071, jul. 2006.

ETHEREUM. **Ethereum**. Disponível em: <<https://www.ethereum.org/>>. Acesso em: 22 jul. 2018.

FELDMAN, R.; DAGAN, I. Knowledge Discovery in Textual Databases (KDT) data structure: The concept hierarchy concept distributions. **Science**, p. 112–117, 1995.

FELDMAN, R.; HIRSH, H. Exploiting background information in knowledge discovery from text. **Journal of Intelligent Information Systems**, v. 9, n. 1, p. 83–97, 1997.

FELDMAN, R.; SANGER, J. **The text mining handbook: Advanced approaches in analyzing unstructured data**. Cambridge: Cambridge University Press, 2007.

FELLOWS, I. Wordcloud: Word clouds. **R package version**, v. 2, p. 109, 2012.

FORD, D. Develop your technology strategy. **Long Range Planning**, v. 21, n. 5, p. 85–95, 1988.

GARCIA-VEGA, M. Does technological diversification promote innovation? **Research Policy**, v. 35, n. 2, p. 230–246, mar. 2006.

GATIGNON, H.; XUEREB, J.-M. Strategic orientation of the firm and new product performance. **Journal of Marketing Research**, v. 34, n. 1, p. 77, fev. 1997.

GLASER, F.; BEZZENBERGER, L. Beyond cryptocurrencies: A taxonomy of decentralized consensus systems. **ECIS 2015 Completed Research Papers**, p. 1–18, 2015.

GOOGLE. **Google**. Disponível em: <<https://www.google.com/>>. Acesso em: 22 jul. 2018.

GOOGLE TRENDS. **Google Trends**. Disponível em: <<https://trends.google.com.br/trends/?geo=BR>>. Acesso em: 22 jul. 2018.

GRANSTRAND, O. Towards a theory of the technology-based firm. **Research Policy**, v. 27, n. 5, p. 465–489, set. 1998.

GREENACRE, M.; HASTIE, T. **The geometric interpretation of correspondence analysis**. [s.l.] Academic Press, 1984.

GREENWICH ASSOCIATES. Wall street blockchain investments top \$1 billion annually. **Greenwich Associates**, 23 jun. 2016.

GRIFFIN, A.; PAGE, A. L. PDMA success measurement project: Recommended measures for product development success and failure. **Journal of Product Innovation Management**, v. 13, n. 6, p. 478–496, nov. 1996.

GUO, Y.; LIANG, C. Blockchain application and outlook in the banking industry. **Financial Innovation**, v. 2, n. 1, p. 1–12, 9 dez. 2016.

HAIR, J. F. et al. **Multivariate Data Analysis: A global perspective**. 7. ed. Upper Saddle River: Prentice Hall, 2009.

HALDANE, A. G.; QVIGSTAD, J. F. The evolution of central banks: A practitioner's perspective. In: BORDO, M. D. et al. (Eds.). . **Central Banks at a Crossroads**. Cambridge: Cambridge University Press, 2014. p. 627–672.

HAMIDA, E. BEN et al. Blockchain for enterprise: Overview, opportunities and challenges. **ICWMC 2017: The Thirteenth International Conference on Wireless and Mobile Communications Blockchain**, p. 83–88, 2017.

HEDMAN, J.; KALLING, T. The business model concept: Theoretical underpinnings and empirical illustrations. **European Journal of Information Systems**, v. 12, n. 1, p. 49–59, 19 mar. 2003.

HOLDEN, W.; MOAR, J. **Blockchain enterprise survey**. [s.l.: s.n.]. Disponível em: <<https://www.juniperresearch.com/researchstore/fintech-payments/blockchain->

enterprise-survey/deployments-benefit-attitudes>. Acesso em: 24 jul. 2018.

KABASHKIN, I. Risk modelling of blockchain ecosystem. **Transport and Telecommunication Institute**, v. 8, p. 432, 2013.

KATZ, M. L.; SHAPIRO, C. Systems competition and network effects. **Journal of Economic Perspectives**, v. 8, n. 2, p. 93–115, 18 fev. 1994.

KEWELL, B.; ADAMS, R.; PARRY, G. Blockchain for good? **Strategic Change**, v. 26, n. 5, p. 429–437, set. 2017.

LAMPORT, L.; SHOSTAK, R.; PEASE, M. The byzantine generals problem. **ACM Transactions on Programming Languages and Systems**, v. 4, n. 3, p. 382–401, 1 jul. 1982.

LEAL, L. M. DAS N. **Text mining**. Porto: Instituto Superior de Engenharia do Porto, 2010.

LEWIS, D. D. et al. RCV 1: A new benchmark collection for text categorization research. **Journal of Machine Learning Research**, v. 5, p. 361–397, 2004.

LINDMAN, J.; TUUNAINEN, V. K.; ROSSI, M. Opportunities and risks of blockchain technologies: A research agenda. **PLOS ONE**, v. 11, n. 10, p. 1533–1542, 3 out. 2017.

LIU, B. Sentiment analysis and subjectivity. In: INDURKHAYA, N.; DAMERAU, F. J. (Eds.). **Handbook of Natural Language Processing**. 2. ed. New York: CRC Press, 2010. p. 627–666.

LIU, B. Sentiment analysis and opinion mining. **Synthesis Lectures on Human Language Technologies**, v. 5, n. 1, p. 1–167, 2012.

LOH, S. **Abordagem baseada em conceitos para descoberta de conhecimento em textos**. Tese (Doutorado em Ciência da Computação)—Porto Alegre: Universidade Federal do Rio Grande do Sul, 2001.

LOH, S. **Listas de stopwords - stoplist (portugues, inglês, espanhol)**. Disponível em: <<http://miningtext.blogspot.com/2008/11/listas-de-stopwords-stoplist-portugues.html>>. Acesso em: 22 jul. 2018.

LOURENÇO, E. B. **Avaliação**: Contribuição da análise de correspondência para a avaliação docent. Dissertação (Mestrado em Educação)—Campinas: Universidade Estadual de Campinas, 1997.

LUÍS, J.; ROSA, G. O significado da palavra para o processamento de linguagem natural. Estudos Lingüísticos XXVII. **Anais...**São José do Rio Preto: 1998

MAINELLI, M.; SMITH, M. Sharing ledgers for sharing economies: An exploration of mutual distributed ledgers (aka blockchain technology). **The Journal of Financial Perspective**, v. 3, n. 3, p. 38–6, 2015.

MANNING, C. D.; RAGHAVAN, P.; SCHÜTZE, H. **Introduction to information**

retrieval. Cambridge: Cambridge University Press, 2007.

MILES, R. E. et al. Organizational strategy, structure, and process. **Academy of Management Review**, v. 3, n. 3, p. 546–562, jul. 1978.

MILGROM, P. R.; ROBERTS, J. D. **Economics, organization and management**. Englewood Cliffs: Prentice-Hall, 1992.

MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada: Uma abordagem aplicada**. Belo Horizonte: Editora UFMG, 2005.

MIRANDA, A. R. DE O. **Descoberta de conhecimento em texto aplicada a um sistema de atendimento aos usuários de um plano de assistência à saúde**. Dissertação (Mestrado em Engenharia Civil)—Rio de Janeiro: Universidade Federal do Rio de Janeiro, 2009.

MOTTA, R. A busca da competitividade nas empresas. **Revista de Administração de Empresas**, v. 35, n. 2, p. 12–16, abr. 1995.

NAKAMOTO, S. **Bitcoin: A peer-to-peer electronic cash system**. Disponível em: <<https://bitcoin.org/bitcoin.pdf>>. Acesso em: 24 jul. 2018.

OCTOPARSE. **Web scraping tool and free Web Crawlers for data extraction**. Disponível em: <<https://www.octoparse.com/>>. Acesso em: 22 jul. 2018.

OH, S.; WALLSTEN, S. Is blockchain hype, revolution, or both? What we need to know. **Technology Policy Institute**, 2018.

ONDRUS, J.; PIGNEUR, Y. Near field communication: an assessment for future payment systems. **Information Systems and e-Business Management**, v. 7, n. 3, p. 347–361, 17 jun. 2009.

PANIKKAR, S. **Adept: An IoT practitioner perspective**. Disponível em: <https://archive.org/details/pdfy-esMcC00dKmdo53-_>. Acesso em: 24 jul. 2018.

PISA, M.; JUDEN, M. Development: Hype vs. reality. **Center for Global Development Policy Pape**, v. 107, 2017.

PLANSKY, J.; O'DONNELL, T.; RICHARDS, K. A Strategist's Guide to Blockchain. **SPRING**, n. 82, 2016.

PORTER, M. E. **Competitive advantage**. New York: Free Press, 1985.

PORTER, M. E. **Vantagem competitiva: Criando e sustentando um desempenho superior**. Rio de Janeiro: Campus, 1992.

RASKIN, M.; YERMACK, D. Digital Currencies, Decentralized Ledgers, and the Future of Central Banking. **National Bureau of Economic Research**, 2016.

ROBLEH, A. et al. Innovations in payment technologies and the emergence of digital currencies. **Bank of England Quarterly Bulletin**, v. Q3, n. 3, p. 262–276, 2014.

ROGERS, E. **Diffusion of innovations**. New York: Free Press, 2003.

SHANBHOGUE, N. M. **Bank of England Quarterly Bulletin**. [s.l.: s.n.].

SHAW, N. The mediating influence of trust in the adoption of the mobile wallet. **Journal of Retailing and Consumer Services**, v. 21, n. 4, p. 449–459, jul. 2014.

SILVA, A. A. S. DA. **Aiuri: Um portal para mineração de textos integrado a grids computacionais**. Dissertação (Mestrado em Engenharia)—Rio de Janeiro: Universidade Federal do Rio de Janeiro, 2007.

SILVA, K. A. DA et al. A utilização do método Cross-Industry Standard Process for Data Mining no processo de mineração de textos: Extração de termos para criação de uma tecnologia assistiva para o auxílio à alunos com deficiência motora. XII Escola Regional de Informática de Banco de Dados. **Anais...Londrina**: 2016

SIMONITE, T. Web pioneer tries to incubate a second digital revolution. **MIT Technology Review**, nov. 2016.

SOARES, F. DE A. **Mineração de textos na coleta inteligente de dados na Web**. Dissertação (Mestrado em Engenharia Elétrica)—Rio de Janeiro: Pontifícia Universidade Católica do Rio de Janeiro, 2008.

TAN, A.-H. Text Mining: The state of the art and the challenges. **Proceedings of the PAKDD 1999 Workshop on Knowledge Discovery from Advanced Databases**, v. 8, p. 65–70, 1999.

TAPSCOTT, D.; TAPSCOTT, A. **Blockchain revolution: How the technology behind bitcoin is changing money, business, and the world**. Kindle Edi ed. [s.l.] Penguin Publishing Group, 2016.

TEECE, D. J. Reflections on “profiting from innovation”. **Research Policy**, v. 35, n. 8, p. 1131–1146, out. 2006.

WALSH, S. T.; LINTON, J. D. The Strategy-Technology Firm Fit Audit: A guide to opportunity assessment and selection. **Technological Forecasting and Social Change**, v. 78, n. 2, p. 199–216, fev. 2011.

WRIGHT, A.; FILIPPI, P. DE. Decentralized blockchain technology and the rise of lex cryptographia. **SSRN**, 2015.

YLI-HUUMO, J. et al. Where is current research on blockchain technology? A systematic review. **PLOS ONE**, v. 11, n. 10, p. 1–27, 3 out. 2016.

ANEXO A – Código Javascript do Google SERPs Extractor

```

1 javascript:(function(){
2
3 output='&lt;html&gt;&lt;head&gt;&lt;title&gt;SEO SERP Extraction Tool&lt;/title&gt;&lt;style type=\`text/css\`&gt;body,table{font-family:Tahoma,Verdana,Segoe,sans-serif;font-size:11px;c
4
5 output+='\`&quot;&lt;table&gt;&lt;tbody&gt;&lt;tr&gt;&lt;td&gt;&lt;a href=\`https://www.chrisains.com/\`&gt;&lt;img src=\`https://www.chrisains.com/wp-content/uploads/2015/06/chrisains.com-logo1
6
7 pageAnchors=document.getElementsByTagName('a');
8 divClasses=document.getElementsByTagName('div');
9 var%20linkcount=0;var%20linkLocation='';
10 var%20linkAnchorText='';
11
12 output+='\`&lt;table&gt;&lt;thead&gt;&lt;tr&gt;&lt;th&gt;ID&lt;/th&gt;&lt;th&gt;Link&lt;/th&gt;&lt;th&gt;Anchor&lt;/th&gt;';
13
14 for(i=0;i&lt;pageAnchors.length;i++){
15
16     if(pageAnchors[i].parentNode.parentNode.getAttribute('class')!='_Rm'){
17
18         var%20anchorText%20=%20pageAnchors[i].textContent;
19         var%20anchorLink%20=%20pageAnchors[i].href;
20         var%20linkAnchor%20=%20anchorLink%20+%20'\`t'+anchorText;
21         var%20anchorID%20=%20pageAnchors[i].id;
22
23         if(anchorLink!=''){
24             if(anchorLink.match(/^(?!google\.|cache|blogger\.com|\.yahoo\.|youtube\.com\/\?gl=|youtube\.com\/results|javascript:|api\.technorati\.com|botw\.org\/search|del\.icio\.us\/ur
25                 if(anchorID.match(/^(?!hdtb_more|hdtb_tsluh_hl).)*$/i)){
26                 linkLocation+=anchorLink+'&lt;br%20/&gt;';
27                 linkAnchorText+=anchorText+'&lt;br%20/&gt;';
28                 linkcount++;
29                 if%20(anchorText%20===%20undefined)%20anchorText%20=%20pageAnchors[i].innerText;output+='\`&lt;tr&gt;';
30                 output+='\`&lt;td&gt;'+linkcount+'&lt;/td&gt;';
31                 output+='\`&lt;td&gt;'+pageAnchors[i].href+'&lt;/a&gt;&lt;/td&gt;';
32                 output+='\`&lt;td&gt;'+anchorText+'&lt;/td&gt;';
33                 output+='\`&lt;/tr&gt;\`n';
34             }
35         }
36     }
37 }
38 }
39
40 output+='\`&lt;/table&gt;&lt;br/&gt;&lt;h2&gt;URL%20List&lt;/h2&gt;&lt;div&gt;';
41 output+=linkLocation;output+='\`&lt;/div&gt;&lt;br/&gt;&lt;h2&gt;Anchor%20Text%20List&lt;/h2&gt;&lt;div&gt;';
42 output+=linkAnchorText;output+='\`&lt;br/&gt;%C2%A0<br/&gt;&lt;p%20align=center&gt;&lt;a%20href=\`https://www.chrisains.com/\`&gt;www.chrisains.com</a>&lt;/p>';
43
44 with(window.open()){document.write(output);document.close();}})();

```


APÊNDICE A – Planilha de URLs do ano de 2015

A	B	C
ID	Link	Anchor
2	http://allcoinsnews.com/2015/11/16/nick-szabo-on-the-history-of-blockchain/	Nick Szabo on the History of Blockchain
3	http://allcoinsnews.com/2015/11/27/with-multichain-create-a-private-blockchain-in-90-seconds/	
4	http://bravenewcoin.com/news/bitalias-is-a-dns-system-for-bitcoin-addresses-on-the-blockchain	
5	http://business.financialpost.com/news/fp-street/how-blockchain-the-technology-behind-bitcoin-could-turn-traditional-banking-on-its-head	How blockchain, the technology behind Bitcoin, could turn traditional ...
7	http://dataconomy.com/2015/09/blockchain-a-new-economic-model/	
8	http://dataconomy.com/2015/10/crypto-securities-long-blockchain-short-wall-street/	
9	http://dataconomy.com/2015/10/wtf-is-the-blockchain-a-guide-for-total-beginners/	
10	http://fortune.com/2015/03/06/bitcoin-book-boom/	The bitcoin book boom
11	http://fortune.com/2015/05/29/blockchain-big-because-scale/	
12	http://fortune.com/2015/07/14/reid-hoffman-is-betting-on-the-block-chain-to-be-as-big-as-the-internet/	
13	http://fortune.com/2015/08/18/filament-blockchain-iot/	
14	http://fortune.com/2015/10/23/bitcoin-blockchain-alliance/	The Bitcoin industry just got its own Justice League
15	http://fortune.com/2015/10/27/secondmarket-bitcoin-dcg/	
16	http://fortune.com/2015/11/04/jamie-dimon-virtual-currency-bitcoin/	Jamie Dimon: Virtual Currency Will Be Stopped
17	http://fortune.com/2015/11/06/visa-mastercard-bitcoin/	What Do MasterCard and Visa Think About Bitcoin?
18	http://fortune.com/2015/11/27/blockchain-music/	
19	http://fortune.com/2015/12/17/ibm-blockchain-for-business/	IBM, JP Morgan, and Others Build a New Blockchain For Business
21	http://hplusmagazine.com/2015/08/19/resilience%E2%80%8A-%E2%80%8Aa-fully-voluntary-wealth-sharing-network-based-on-the-blockchain/	
22	http://hplusmagazine.com/2015/11/18/blockchain-technology-smart-contracts-and-smart-property/	
23	http://insidebitcoins.com/news/100-banks-approach-blockchain-info/35791	100 Banks Approach Blockchain.info

APÊNDICE B – Planilha de URLs do ano de 2016

	A	B	C
1	ID	Link	Anchor
2	1059	http://allafrica.com/stories/201612200631.html	Kenyan Government Embraces Blockchain Technology Set to ...
3	1232	http://bigthink.com/robby-berman/bitcoin-schmitcoin-the-real-mind-blower-is-the-blockchain-behind-it	
5	958	http://bravenewcoin.com/news/new-blockchain-e-voting-service-compliments-abu-dhabi-economic-vision/	New blockchain e-Voting service complements Abu Dhabi economic ...
6	74	http://bravenewcoin.com/news/shakeup-at-r3-cev-blockchain-consortium-as-banks-leave/	Shakeup at R3 CEV blockchain consortium as banks leave
7	1196	http://bravenewcoin.com/news/south-korea-plans-national-digital-currency-using-a-blockchain/	
8	786	http://coinjournal.net/keepkey-shapeshift-partnership-cryptocurrency-exchange/	CoinJournal
12	642	http://deloitte.wsj.com/cfo/2016/02/26/beyond-bitcoin-blockchain-is-coming-to-disrupt-your-industry-weekend-reading/	Wall Street Journal
14	1285	http://financialpost.com/entrepreneur/fp-startups/are-we-ready-to-place-our-trust-in-blockchain-when-it-comes-to-our-money-and-security	
15	410	http://fortune.com/2016/02/18/blythe-masters-blockchain/	
16	5	http://fortune.com/2016/03/04/crisis-in-bitcoin-rise-of-blockchain/	The Crisis in Bitcoin and the Rise of Blockchain
17	810	http://fortune.com/2016/03/30/blockchain-lehman-brothers/	
18	52	http://fortune.com/2016/04/29/ibm-blockchain-cloud-business-build/	IBM: Build a Blockchain in Our Cloud
19	380	http://fortune.com/2016/05/04/larry-summers-bitcoin/	Larry Summers Believes in Blockchain, Not Sure About Bitcoin
20	130	http://fortune.com/2016/05/08/why-blockchains-will-change-the-world/	
21	1201	http://fortune.com/2016/05/15/blockchain-reinvents-power-grid/	
22	421	http://fortune.com/2016/05/15/leaderless-blockchain-vc-fund/	
23	262	http://fortune.com/2016/05/23/blockchain-definition/	
24	705	http://fortune.com/2016/06/23/ripple-blockchain-banks/	
25	282	http://fortune.com/2016/06/28/ibm-blockchain/	
26	432	http://fortune.com/2016/07/07/blockchain-r3/	How Banks Will Stop Snoops From Using the Blockchain to Front ...

2016-1 até 31-12-2016 Ordenada

APÊNDICE C – Planilha de URLs do ano de 2017

	A	B	C
1	ID	Link	Anchor
2	1512	http://algononline.com/industry/ibm-inks-blockchain-alliance-with-10-colossal-food-giants-multi-firm-collaboration-aims-to-regularize-global-food-safety	IBM inks blockchain alliance with 10 colossal food giants, multi-firm ...
3	1354	http://bernews.com/2017/11/video-premier-blockchain-cryptocurrency/	
4	1169	http://bitcoinist.com/blockchain-ad-network-adex-crosses-real-sector-ink-partnership/	
5	1219	http://bitcoinist.com/chinas-first-blockchain-platform-antshares-rebrands-neo/	
6	1446	http://bitcoinist.com/matrix-safer-faster-flexible-intelligent-blockchain/	
7	922	http://bitcoinist.com/pentagon-drafts-blockchain-technology-security-purposes/	
8	1389	http://bitcoinist.com/pundi-x-integrates-nem-blockchain/	
9	1399	http://blocktribune.com/reality-shares-nasdaq-create-blockchain-economy-index/	BlockTribune
10	1143	http://business.financialpost.com/investing/pot-stocks-yesterdays-news-as-canada-gripped-by-bitcoin-fever	
11	363	http://business.financialpost.com/news/fp-street/mogo-shares-surge-on-plan-to-introduce-blockchain-technology	Financial Post
12	1576	http://business.financialpost.com/pmn/business-pmn/legal-weed-hurts-prognosis-for-california-medical-pot-market	
13	986	http://bwdisrupt.businessworld.in/article/How-Andhra-Pradesh-Is-Creating-An-Ecosystem-For-Fintech-Startups/11-08-2017-123895/	BW Businessworld
24	497	http://fortune.com/2017/02/09/accnture-blockchan-security/	
25	145	http://fortune.com/2017/02/28/ethereum-jpmorgan-microsoft-alliance/	Big Business Giants From Microsoft to JP Morgan Are Getting ...
26	324	http://fortune.com/2017/05/23/blockchain-chasm-of-death-bitcoin/	
27	181	http://fortune.com/2017/06/22/google-blockchain-funding-series-b/	
28	229	http://fortune.com/2017/07/11/ibm-blockchain-hyperledger-fabric/	
29	525	http://fortune.com/2017/07/19/bitcoin-ethereum-blockchain-sofi/	Why Ethereum Is Much More Valuable Than Bitcoin: SoFi CEO
30	92	http://fortune.com/2017/07/26/sec-icos/	The SEC's Big Digital Coin Ruling: What It Means
31	216	http://fortune.com/2017/07/28/bitcoin-cryptocurrency-ico/	Cryptocurrency ICOs Are Making Bitcoin Startups Richer than VCs ...

2017-1 até 31-12-2017 Ordenada