

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

LUIZ HERMES SVOBODA JUNIOR

**Utilizando Blockchain para criação de uma
plataforma para escrita colaborativa de
artigos científicos**

Plano de Estudos e Pesquisa

Orientador: Prof. Dr. José Palazzo Moreira de
Oliveira

Porto Alegre
2017

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Rui Vicente Oppermann

Vice-Reitora: Prof^a. Jane Fraga Tutikian

Pró-Reitor de Pós-Graduação: Prof. Celso Giannetti Loureiro Chaves

Diretora do Instituto de Informática: Prof^a. Carla Maria Dal Sasso Freitas

Coordenador do PPGC: Prof. João Luiz Dihl Comba

Bibliotecária-chefe do Instituto de Informática: Beatriz Regina Bastos Haro

RESUMO

A liberdade para criação e acesso à informação evoluiu drasticamente com o advento da Web. Embora essa evolução possa ser observada em todas as áreas de conhecimento, o acesso à parte do conteúdo científico ainda está sob controle de algumas organizações. Uma reação a esse controle são as iniciativas *open access* que fornecem acesso livre ao conteúdo publicado. A opção por publicações abertas levanta a discussão sobre como manter os direitos autorais sobre o trabalho, como efetuar um processo de revisão aberto e justo, e como recompensar os autores pelo seu trabalho, baseado no nível de colaboração de cada indivíduo. Desta forma, este trabalho tem como objetivo o desenvolvimento de uma plataforma aberta e colaborativa para escrita e revisão de trabalhos científicos, utilizando como base para isso a tecnologia de *blockchain*. O uso de *blockchain* para este fim traz consigo os benefícios inerentes a sua utilização, tais como durabilidade, confiabilidade e longevidade, já que os dados inseridos na *chain* não podem ser alterados. Outro ponto importante da utilização desta tecnologia é o seu caráter de descentralização, o que facilita a contribuição entre os participantes e evita o controle sobre as informações armazenadas.

Palavras-chave: Publicação científica, *open access*, *peer review*, *blockchain*.

SUMÁRIO

1 INTRODUÇÃO E MOTIVAÇÃO	5
2 OBJETIVOS E RESULTADOS ESPERADOS.....	8
3 METODOLOGIA	10
4 CRONOGRAMA DE ATIVIDADES	11
REFERÊNCIAS.....	12

1 INTRODUÇÃO E MOTIVAÇÃO

A quantidade de informação disponível vem crescendo exponencialmente ao longo dos últimos anos. Este crescimento afeta todas as áreas de conhecimento, desde informações livremente publicadas em meios como a Web, até trabalhos acadêmicos os quais são submetidos a processos de revisão antes de se tornarem públicos. Neste contexto, destaca-se o aumento da quantidade de trabalhos científicos. Utilizando como base o banco de dados de publicações da área de computação, o DBLP,¹, verificamos um crescimento constante de aproximadamente 3% ao longo dos últimos 5 anos. Em (LAWRENCE, 2001) o autor comenta que a quantidade de conteúdo científico de longe excede a capacidade dos pesquisadores de consumi-lo. Além disso, neste mesmo trabalho é ressaltada a importância do acesso livre ao conteúdo publicado, ou *open access publications*.

Existe um interesse crescente da comunidade acadêmica em publicações livremente distribuídas. Esse formato de publicação compreende a disponibilização do conteúdo da pesquisa de maneira livre, principalmente através de meios eletrônicos. A busca por publicações abertas se dá principalmente pelo alcance que podem ter e pela velocidade com que o conteúdo está disponível para o público. Em (HARNAD; BRODY, 2004) é feita uma comparação do impacto, considerado como sendo o número e citações, entre publicações abertas em relação ao modelo de publicação tradicional. Os autores identificaram uma correlação forte entre o número de citações de um trabalho com o fato de eles estar disponível abertamente para consulta.

Desta forma, em (LAWRENCE, 2001), o autor recomenda como estratégia para maximizar o impacto e acelerar a evolução da Ciência, que autores e organizações facilitem o acesso ao conteúdo científico. Outros estudos sobre o mesmo assunto mostram que após determinado período a partir da data da publicação da pesquisa de forma aberta, mais da metade dos pesquisadores estariam dispostos a pagar pelo conteúdo consumido (WALKER, 2002). Para evidenciar o impacto de publicações abertas, o estudo conduzido em (ANTELMAN, 2004) verifica que mesmo em áreas de pesquisa como Filosofia, cuja adoção por publicações em meios eletrônicos ocorreu tardiamente em comparação a outras, o número de citações de conteúdo distribuído livremente é 45% maior em relação a trabalhos disponibilizados da forma tradicional. A partir dessa análise, os autores concluem que a hipótese inicialmente levantada em (LAWRENCE, 2001) para Computação se aplica também a outras áreas de pesquisa.

¹<http://dblp.uni-trier.de/statistics/publicationsperyear.html>

Embora diversas iniciativas de criação de *journals* e repositórios para acesso livre a conteúdo científico tenham sido criados, poucos meios relevantes continuam ativos até a atualidade. Como exemplo, podemos citar arXiv², Public Library of Science (PLOS)³, BioMed⁴ e Directory of Open Access Journals (DOAJ)⁵, cada um com suas características para submissão, revisão e disponibilização de conteúdo. Em (BJöRK, 2004) o autor faz uma avaliação sobre 6 aspectos relacionadas a publicações abertas e como elas seriam abordadas em três cenários distintos de publicação: *open access journals*, repositórios abertos e páginas pessoais dos seus autores.

Outro fator que afeta o impacto das publicações é a colaboração entre diferentes pesquisadores. De acordo com (KATZ; MARTIN, 1997), um colaborador poderia ser definido como alguém que fornece alguma entrada ou contribui com parte da pesquisa. Outra definição seria a de uma pessoa que contribui diretamente para os resultados da pesquisa. Conforme podemos observar, a definição de colaborador não é precisa, mas sim uma convenção do grupo ou área em que os pesquisadores atuam.

Uma das formas de avaliar a colaboração de um trabalho é através dos seus autores (MEADOWS; CONNOR, 1971). Embora existam críticas em relação a esse método para avaliação de colaboração, já que em alguns casos autores poderiam ser incluídos apenas por razões sociais e não deveriam receber crédito pelo trabalho (HUTH, 1993), verificou-se que o nível de colaboração influencia o impacto do trabalho na comunidade, se considerarmos que trabalhos com múltiplos autores acabam sendo mais citados (LAWANI, 1986).

Uma das principais motivações para a colaboração é o fato de que a pesquisa moderna demanda conhecimento em diversas áreas. Com frequência, apenas um indivíduo não possui conhecimento suficiente em todas elas. Assim, a opção de colaboração com outros pesquisadores que já possuem tal conhecimento é uma alternativa. Outro fato interessante da colaboração é a diferença de perspectivas e ideias sobre o trabalho, características que enriquecem as discussões e o entendimento sobre os resultados. Outro aspecto da colaboração é que ela pode ser utilizada para medir a qualidade do pesquisador, grupo de pesquisa ou instituição. Como exemplo deste uso, podemos citar a avaliação feita pela CAPES dos programas de pós-graduação brasileiros⁶.

A opção pela publicação de conteúdo científico de maneira aberta e criado de

²<https://arxiv.org/>

³<https://www.plos.org/>

⁴<https://www.biomedcentral.com/>

⁵<https://doaj.org/>

⁶<http://capes.gov.br/avaliacao/sobre-a-avaliacao>

maneira colaborativa, traz à tona discussões sobre garantia de direitos autorais sobre a pesquisa, processo de revisão (*peer review*) aberto e modelos de recompensa aos autores pelo trabalho. Esses e outros aspectos são discutidos em (BJöRK, 2004). Por fim, conforme mencionado anteriormente, com a colaboração entre diversos autores para produção de conteúdo, é necessário um método para determinar o nível de contribuição de cada indivíduo no trabalho.

Este trabalho tem como objetivo a concepção de uma proposta de plataforma aberta e colaborativa para a elaboração e revisão de trabalhos científicos utilizando como base a tecnologia de *blockchain*. Conforme (Dinh et al., 2017), *blockchain* basicamente consiste em uma rede de nodos na qual não há necessidade de confiança entre todos os participantes, assim como não existe um nodo centralizador ou que exerça qualquer tipo de controle sobre o a rede. Os nodos mantêm um conjunto de estados compartilhado e executam transações que alteram e validam esses estados. Assim, pode-se afirmar que *blockchain* é um tipo de estrutura que armazena o histórico de estados e suas mudanças, na qual todos os nodos do sistema concordam com as transações e ordem em que ocorrem. Estas características garantem a esta estrutura as características de integridade, durabilidade e longevidade, já que enquanto a maior parte do poder computacional da rede não estiver sendo utilizada para violá-la, há garantia de que os dados não serão alterados nem perdidos (NAKAMOTO, 2008).

Por padrão, todas as implementações de *blockchain* possuem o conceito de *smart contract* (SZABO, 2017). No caso de uso mais famoso de *blockchain*, o Bitcoin⁷, esse contrato contém o conjunto de cláusulas que descreve como uma transação financeira deve ocorrer. Neste caso, o conceito de transação é definido como sendo um acordo financeiro em que ocorre a troca de valor entre os participantes da rede. Mais recentemente, surgiram implementações de *blockchain* que permitem aos seus participantes definirem seus próprios contratos, de forma que eles sejam executados na rede quando uma transação é iniciada.

Através da utilização das tecnologias de *blockchain* e contratos inteligentes, esta dissertação fará uma proposta de plataforma que ofereça as garantias necessárias para que pesquisadores possam desenvolver trabalhos colaborativos, assegurando as garantias de autoria e níveis de contribuição. Além disso, essa plataforma poderá prover suporte a uma modelo de revisão aberto onde os revisores são conhecidos e podem ser inclusive recompensados pelas suas revisões.

⁷<https://bitcoin.org>

2 OBJETIVOS E RESULTADOS ESPERADOS

O objetivo deste trabalho é a elaboração de uma proposta de plataforma aberta e colaborativa para criação e revisão aberta de artigos científicos. Mais especificamente, a definição de contratos inteligentes que descrevam as interações necessárias contemplando ações de elaboração, revisão e acesso aos artigos desenvolvidos, bem como a especificação de um modelo de negócio.

Conforme mencionado na seção anterior, a utilização da tecnologia de *blockchain* para a implementação da ferramenta, visa agregar à plataforma algumas das suas características como segurança e durabilidade. Para alcançar tal objetivo, este trabalho fará utilização do conceito de *smart contracts* disponível em plataformas como Ethereum¹ e Hyperledger² para a descrição das cláusulas que irão definir as interações entre os participantes.

Dentre as interações que devem estar disponíveis para utilização pelos autores e revisores, podemos destacar (1) a coleta de interessados em colaborar com o desenvolvimento de uma pesquisa em determinada área, (2) garantir que os envolvidos sejam identificados e que o direito sobre o resultado do seu trabalho seja mantido, (3) propor um processo de revisão aberta no qual revisores sejam alocados de acordo com seu nível de conhecimento, (4) garantir o acesso livre ao conteúdo produzido e finalmente (5) especificar um modelo de negócio que escale conforme a quantidade de participantes na rede.

Conforme citado anteriormente neste trabalho, os itens que estão listados como seu objetivo endereçam cada uma das características identificadas como positivas no desenvolvimento de trabalhos científicos, como por exemplo, estímulo a colaboração entre autores e a liberdade de acesso ao conteúdo produzido. Outros aspectos discutidos em (BJÖRK, 2004) também devem ser abordados, sendo o principal deles e também discutido em (PöSCHL, 2010) a definição de um modelo de negócio capaz de garantir a manutenção da rede, além de escalar conforme necessidade. Sob este aspecto pode-se avaliar alternativas como a adotada no PLOS em que os autores (ou instituições) arcam com os custos de publicação ou a do arXiv na qual foi identificado um conjunto de instituições parceiras, as quais financiam a manutenção do sistema.

Outra característica a ser avaliada com cuidado diz respeito à recompensa dos autores pelas publicações. Seria bastante ingênuo acreditar que autores utilizariam a pla-

¹<https://www.ethereum.org>

²<https://www.hyperledger.org>

taforma apenas por suas características de segurança, acesso livre ao conteúdo, etc. Além disso, autores estão preocupados com a reputação do meio no qual buscam efetuar a publicação. Assim, é importante assegurar que o trabalho submetido pela plataforma seja de alguma forma recompensado em caso de publicação, uma vez que deve ser submetido a um criterioso processo de revisão. O mesmo pensamento se aplica aos revisores. A opção por um processo aberto de revisão se dá pelo benefício da transparência (PöSCHL, 2010). Este modelo de revisão é utilizado pelo BioMed e se provou positivo já que a identificação dos revisores torna possível que estes sejam cobrados pelo resultado, assim como também são elegíveis a incentivos.

Como resultado deste trabalho, é esperada a implementação de uma plataforma aberta e colaborativa para a escrita de artigos científicos, onde os pares podem colaborar de acordo com sua área de interesse, submetendo seus trabalhos a uma processo de revisão aberta. Essa plataforma adota princípios comprovadamente eficientes para o aumento do impacto dos trabalhos publicados, já que incentiva colaboração e acesso livre ao conteúdo, mantendo o direito sobre o resultado aos autores. Além disso, deverá ser incluída uma forma de verificação no nível de contribuição de cada autor, a fim de evitar que eles sejam incluídos nos trabalhos por quaisquer outras razões (HUTH, 1993) que não em virtude da sua contribuição.

3 METODOLOGIA

A metodologia de trabalho prevista compreende uma análise bibliográfica extensiva sobre os requisitos para adoção de um padrão de livre acesso à conteúdo científico. O objetivo desta etapa, além de verificar as vantagens desse modelo, é garantir que autores e revisores tenham seus direitos mantidos, evitando assim possíveis penalidades ou perdas de garantias para ambos. Em seguida, avaliaremos métodos para execução de revisão aberta (*open peer review*), conforme citado em (PöSCHL, 2010). Da mesma forma como no estudo anterior, o objetivo é entender os modelos possíveis para implementação de um modelo de revisão aberta e quais as vantagens e desvantagens de cada abordagem.

Uma vez que os conceitos fundamentais tenham sido entendidos, será iniciada a avaliação das plataformas de *blockchain* disponíveis para utilização. Nesta etapa será realizada uma avaliação para verificação das suas funcionalidades. De modo concomitante, será conduzida uma avaliação do suporte à criação de *smart contracts* em cada plataforma. Essa etapa é fundamental uma vez que é necessário garantir que a linguagem a ser utilizada para definição do contrato tenha expressividade suficiente para descrever todas as cláusulas necessárias. Por fim, em virtude dos custos associados à utilização de *blockchain* para armazenamento de dados, avaliaremos opções fora da *chain* para persistência de dados. Nesta etapa avaliaremos ferramentas como IPFS¹, por exemplo.

A próxima etapa compreende a implementação de um protótipo cujo objetivo é fornecer suporte ao processo de coleta de colaboradores e desenvolvimento do trabalho de maneira colaborativa entre os participantes. Considerando que os envolvidos no desenvolvimento do trabalho tenham finalizado seu esforço inicial, deve ser possível efetuar a submissão para revisão. Nesta fase é importante verificar que os revisores sejam alocados de forma adequada para que o resultado da revisão agregue e garanta mais qualidade ao trabalho final. É importante que o protótipo da plataforma atenda aos requisitos citados na seção anterior deste documento. Esta fase também compreende o suporte ao acesso livre ao conteúdo, bem como a implementação de um modelo de negócios sustentável para garantir a evolução da plataforma, além de recompensar os envolvidos em todo o processo.

Por fim, após as etapas de criação do protótipo e avaliação, será desenvolvida a redação da dissertação e de um artigo científico. O meio de publicação do artigo (conferência ou *journal*) deverá ser definido ao longo do desenvolvimento da dissertação.

¹<https://ipfs.io/>

REFERÊNCIAS

- ANTELMAN, K. Do open-access articles have a greater research impact? **College Research Libraries**, v. 65, n. 5, p. 372–382, 2004. ISSN 2150-6701. Disponível em: <<http://crl.acrl.org/index.php/crl/article/view/15683>>.
- BJÖRK, B.-C. Open access to scientific publications – an analysis of the barriers to change? **Information Research**, v. 9, 2004.
- Dinh, T. T. A. et al. Untangling Blockchain: A Data Processing View of Blockchain Systems. **ArXiv e-prints**, ago. 2017.
- HARNAD, S.; BRODY, T. Comparing the impact of open access (oa) vs. non-oa articles in the same journals. **D-Lib Magazine**, v. 10, n. 6, June 2004. Disponível em: <<https://eprints.soton.ac.uk/260207/>>.
- HUTH, E. J. Stealing into print: Fraud, plagiarism, and misconduct in scientific publishing. **Publishing Research Quarterly**, v. 9, n. 2, p. 78–79, Jun 1993. ISSN 1936-4792. Disponível em: <<https://doi.org/10.1007/BF02680404>>.
- KATZ, J.; MARTIN, B. R. What is research collaboration? **Research Policy**, v. 26, n. 1, p. 1 – 18, 1997. ISSN 0048-7333. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0048733396009171>>.
- LAWANI, S. M. Some bibliometric correlates of quality in scientific research. **Scientometrics**, v. 9, n. 1, p. 13–25, Jan 1986. ISSN 1588-2861. Disponível em: <<https://doi.org/10.1007/BF02016604>>.
- LAWRENCE, S. Online or invisible? **Nature**, v. 411, p. 521, 2001.
- MEADOWS, A.; CONNOR, J. O. Bibliographical statistics as a guide to growth points in science. **Science Studies**, v. 1, n. 1, p. 95–99, 1971. Disponível em: <<https://doi.org/10.1177/030631277100100107>>.
- NAKAMOTO, S. **Bitcoin: A peer-to-peer electronic cash system**. 2008. Disponível em: <<http://bitcoin.org/bitcoin.pdf>>.
- PöSCHL, U. Interactive open access publishing and public peer review: The effectiveness of transparency and self-regulation in scientific quality assurance. **IFLA Journal**, v. 36, n. 1, p. 40–46, 2010. Disponível em: <<https://doi.org/10.1177/0340035209359573>>.
- SZABO, N. **Smart Contracts: Building Blocks for Digital Market**. 2017. Disponível em: <http://www.fon.hum.uva.nl/rob/Courses/InformationInSpeech/CDROM/Literature/LOTwinterschool2006/szabo.best.vwh.net/smart_contracts_2.html>.
- WALKER, T. J. **Two societies show how to profit by providing free access. Learned Publishing (forthcoming)**. [Online]. Available: <http://csssrvr.entnem.ufl.edu/walker/epub/ALPSPmsDS2.pdf>. 2002.