

Ricardo César Gonçalves Sant'Ana

Moisés Lima Dutra

Guilherme Ataíde Dias

Organizadores

WIDaT 2018

II WORKSHOP DE INFORMAÇÃO,
DADOS E TECNOLOGIA

ANAIS
WIDaT 2018

Organização do WIDaT 2018

- **Organização Geral:**

Guilherme Ataíde Dias (PPGCI-UFPB) - Coordenador geral do evento
Moisés Lima Dutra (PPGCIN-UFSC) - Vice-coordenador

- **Coordenador da Comissão Científica:**

Ricardo César Gonçalves Sant'Ana (PPGCI-UNESP)

- **Comissão científica**

Adilson Luiz Pinto (PPGCIN-UFSC)
Ana Alice Baptista (Universidade do Minho, Portugal)
Ana Carolina Simionato (PPGCI-UFSCar)
Angela Maria Grossi de Carvalho (PPGCI-UNESP)
Bernardina Maria Juvenal Freire de Oliveira (PPGCI-UFPB)
Cristian Berrío-Zapata (PPGCI-UFPA)
Dalton Lopes Martins (FCI-UnB)
Denysson Axel Ribeiro Mota (PPGB-UFCA)
Douglas Dyllon Jeronimo de Macedo (PPGCIN-UFSC)
Ed Porto Bezerra (PPGI-UFPB)
Edgar Bisset Alvarez (PPGCIN-UFSC)
Edna Gusmão de Goés Brennand (MPGOA-UFPB)
Edna Gomes Pinheiro (DCI-UFPB)
Elaine Parra Affonso (FATEC-SP)
Elvis Fusco (UNIVEM-Marília)
Enrique Muriel Torrado (PPGCIN-UFSC)
Evandro de Barros Costa (IC-UFAL)
Fábio Paraguaçu (IC-UFAL)
Fernando de Assis Rodrigues (PPGCI-UNESP)
Gustavo Medeiros de Araújo (PPGCIN-UFSC)
Henry Pôncio Cruz de Oliveira (PPGCI-UFPB)
Joana Coeli Ribeiro Garcia (PPGCI-UFPB)
José Eduardo Santarém Segundo (USP-FFCLRP)
Leonardo Castro Botega (UNIVEM-Marília)
Luana Farias Sales Marques (PPGCI-IBICT-UFRJ)
Marckson Roberto Ferreira de Sousa (PPGCI-UFPB)
Luís Fernando Sayão (CNEN)
Marcelo Morandini (EACH-USP)
Márcio Matias (PPGCIN-UFSC)
Marcos Mucheroni (CBD-USP)
Marynice de Medeiros Matos Autran (PPGCI-UFPB)

Maurício Barcellos Almeida (PPGGOC-UFMG)
Moisés Lima Dutra (PPGCIN-UFSC)
Plácida Leopoldina V. da Costa Santos (PPGCI-UNESP)
Pedro Luiz Pizzigatti Corrêa (POLI-USP)
Renata Baracho (PPGGOC-UFMG)
Ricardo César Gonçalves Sant'Ana (PPGCI-UNESP)
Robson Rodrigues Lemos (UFSC-Araranguá)
Rogério Ramalho (PPGCI-UFSCar)
Ryan Ribeiro de Azevedo (UFRPE-UAG)
Sandra de Albuquerque Siebra (PPGCI-UFPE)
Sandro Rautenberg (DECOMP-UNICENTRO)
Silvana Aparecida Borsetti G. Vidotti (PPGCI-UNESP)
Virginia Bentes Pinto (PPGCI-UFC)
Wagner Junqueira de Araújo (PPGCI-UFPB)
Zaira Regina Zafalon (PPGCI-UFSCar)

- **Coordenador do Cerimonial:**

André Luiz Dias de França (PPGCI-UFPB)

- **Coordenador da Equipe Técnica Local:**

Laerte Pereira da Silva Júnior (CCHLA-UFPB)

- **Equipe Técnica Local:**

Adriana Alves Rodrigues (PPGCI-UFPB)
Antonio Felipe dos Santos (MPGOA-UFPB)
Débora Gomes de Araújo (PPGCI-UFPB)
Pedro Augusto de Lima Barroso (PPGCI-UFPB)
Pollianna Marys de Souza e Silva (PPGCI-UFPB)
Renata Lemos dos Anjos (PPGCI-UFPB)

PROTEÇÃO AUTORAL DE DADOS:

Uso de Blockchain

AUTHORAL PROTECTION OF DATA:

Use of Blockchain

Elizabete Cristina de Souza Aguiar Monteiro¹, Jacquelin Teresa Camperos Reyes², Elaine Parra Affonso³, Ricardo César Gonçalves Sant'Ana⁴

(1) Universidade Estadual Paulista, Marília, ecsamonteiro@gmail.com

(2) Universidade Estadual Paulista, Marília, jcamperosreyes@gmail.com

(3) FATEC, Presidente Prudente, elainepff@gmail.com

(4) Universidade Estadual Paulista, Marília, ricardo.santana@unesp.br

Resumo:

A tecnologia *blockchain* pode apresentar potencial para proteger o direito autoral de dados quando estes estão envolvidos em ambientes digitais. Com o intuito de explicitar como *blockchain* tem sido abordada para questões de proteção do direito autoral de dados, foi realizada uma revisão sistemática de literatura, junto com análise de conteúdo, em trabalhos publicados na base de dados *Scopus*. Os resultados demonstram que o uso da tecnologia *blockchain* para a proteção de direitos autorais é pouco recorrente nos 80 estudos analisados, apenas oito apresentam aplicação dessa tecnologia para proteção autoral de dados. Entre as aplicações que emergem nesse cenário há plataformas como Mycelia, criação de marca d'água e contratos inteligentes. Conclui-se que o uso de *blockchain* no contexto de proteção de dados autorais não é foco das pesquisas recuperadas, situação que torna essa temática uma oportunidade para pesquisas em diversas áreas do conhecimento.

Palavras-chave: Direitos autorais. Blockchain. Propriedade intelectual. Dados.

Abstract: Blockchain technology may have the potential to protect copyright when data are involved in digital environments. In order to explain how blockchain has been approached for copyright protection issues, a systematic review of the literature, together with content analysis, was carried out in published works in the *Scopus* database. The results show that the use of blockchain technology for copyright protection is not very recurrent in the 80 studies analyzed, only eight present application of this technology for data protection. Among the applications that emerge in this scenario are platforms like Mycelia, watermark creation and smart contracts. It was observed that blockchain is involved with studies and applications in several areas of knowledge, marking this topic as a fertile field for research in the scope of protection and verification of data copyright.

Keywords: Copyright. Blockchain. Intellectual property. Data.

I INTRODUÇÃO

O avanço na ciência e na produção de dados em diversas atividades da sociedade podem proporcionar oportunidades no âmbito da colaboração internacional, geração de conhecimentos e progresso da ciência, porém, apresenta desafios quando relacionados à reutilização e a proteção autoral dos dados.

O compartilhamento ou publicação de dados oferecem vantagens para os cientistas pois, quando os pesquisadores publicam seus dados, o seu trabalho pode ser compartilhado entre as fronteiras regionais, nacionais e internacionais (TENOPIR et al., 2015).

Para o pesquisador que busca melhorar seu *status* profissional e seu reconhecimento perante seus pares (agradecimentos e citações), o compartilhamento de seus dados propicia oportunidades para ser reconhecido como o autor, ser citado por membros mais proeminentes de seu campo de atuação e receber o reconhecimento formal (KIM; STANTON, 2013, 2016). Da mesma forma, a divulgação de dados fora do contexto científico impacta positivamente nas carreiras dos profissionais.

Os fatores apresentados instigam o compartilhamento dos dados, no entanto, os custos na geração ou coleta, no tratamento e na gestão de dados são relativamente dispendiosos e os detentores dos dados têm recorrido a estratégias para proteger a autoria dos dados. Um recurso utilizado para questões de proveniência e garantia da autoridade de dados é a tecnologia *blockchain*.

Segundo Kim e Laskowski (2018) *blockchain* é um registro público baseado em uma arquitetura distribuída, uniformemente replicado em todos os seus nós, à prova de adulteração, que fornece um método para estabelecer a existência de uma transação em um momento específico, podendo ser verificada por qualquer parte interessada, incluindo o rastreamento da proveniência dos dados.

Tecnicamente, *blockchain* é um sistema distribuído público que contém todas as transações que já foram executadas no sistema. No contexto da gestão de dados de pesquisa, *blockchain* permite gerir esses dados de forma descentralizada, com a possibilidade da auto-regulação por meio de uma infraestrutura compartilhada para transações ou fluxos dos dados. A tecnologia também permite que pesquisadores interajam com o conteúdo em qualquer etapa do ciclo de vida dos dados. Ainda, interessante ressaltar que, o uso de *blockchain* nas questões vinculadas à proteção do direito autoral:

[...] permite, por um lado, a criação de um registro permanente dos acordos entre os autores dos dados e os reutilizadores, com a possibilidade de verificar a conformidade a qualquer momento e, por outro lado, um maior nível de granularidade na definição das condições de reutilização. (PANESCU; MANTA, 2018)

Assim, o registro é fator significativo, na medida em que confere a autenticidade do protocolo original e a segurança jurídica ao autor, facilita a comprovação da criação da obra, assegura o direito autoral na publicação e a obtenção de um certificado de registro. Ainda, para Karafiloski (2017) quando se tem o registro de data e hora das transações, é possível obter prova no tribunal, no caso de uma disputa de propriedade.

2 OBJETIVOS

No contexto desta pesquisa busca-se investigar a aplicação da tecnologia *blockchain* para proteção do direito autoral de dados e relatar como está sendo utilizada.

3 PROCEDIMENTOS METODOLÓGICOS

A metodologia deste trabalho consiste em uma revisão sistemática de literatura a fim de verificar como a tecnologia *blockchain* tem sido aplicada para comprovação de direitos autorais em dados. As estratégias utilizadas nesta metodologia foram baseadas no protocolo de revisão sistemática de Kitchenham (2004). Para tanto, realizaram-se as seguintes estratégias:

- a) Base de dados: Foi escolhida a *Scopus* por fornecer um serviço eficaz na obtenção de informações de produção científica (NOBLEJAS; RODRIGUEZ, 2014), envolvida com diversas áreas de conhecimento, visto que a temática deste trabalho abarca questões da Computação, do Direito e da Ciência da Informação;
- b) Termo descritor: A busca foi realizada no dia 2 de outubro de 2018, por meio da seguinte estratégia de busca: (TITLE-ABS-KEY (BLOCKCHAIN) AND (TITLE-ABS-KEY (“intellectual property” OR copyright) AND TITLE-ABS-KEY (blockchain))). Os termos foram escolhidos considerando o foco da pesquisa na tecnologia *blockchain* como opção para proteção do direito autoral de dados;
- c) Tipos de documentos: Foram analisados todos os documentos recuperados compostos por artigos de periódicos, capítulos de livros e artigos publicados em anais de eventos;
- d) Ano de publicação: Foram considerados para a análise todos os documentos recuperados mediante a expressão de busca utilizada, assim, a coleta não foi delimitada por valor temporal;
- e) Critério de exclusão: Não foi excluído nenhum documento recuperado;
- f) Processo de seleção: Após a expressão de busca ser submetida à base de dados *Scopus*, os atributos referentes a cada documento recuperados foram armazenados em *software* de planilhas, extraindo as seguintes informações: Ano de publicação, título do artigo, quantidade de citações, resumo e palavras-chaves;
- g) Processo de análise: Foi aplicada Análise de conteúdo, adotando a análise léxica, tendo como unidade de registro a palavra, e utilizando a regra de enumeração “presença” (BARDIN, 1977). Para tanto, foi utilizado a ferramenta *textalyser*, que possibilita realizar a análise estatística de um texto, permitindo descobrir palavras-chave. Posteriormente utilizou-se a linguagem R para geração da nuvem de palavras;
- h) Mediante a leitura dos resumos realizou-se a análise categorial proposto no método Análise de conteúdo para verificar a principal abordagem no uso de *blockchain* no contexto de proteção de dados autorais. Esta atividade permitiu a construção do quadro no Apêndice A com as seguintes categorias: tipo de dado, técnica/processo e descrição do processo.

Foram utilizados quadros e figuras para representar os dados extraídos dos documentos, com a finalidade de demonstrar a frequência das palavras nos textos e a quantidade de trabalhos por ano. Por meio desses dados é possível inferir sobre o objeto de estudo deste trabalho.

4 RESULTADOS

Foram recuperados 80 trabalhos na base de dados *Scopus* de acordo com a estratégia de busca *blockchain and “copyright or intellectual property” and data*. Após a coleta, os resumos foram agrupados em um único arquivo, resultando em um arquivo com 14 páginas de texto. Este arquivo foi inserido na ferramenta *textalyser* para realizar a análise de contagem de palavras e frequência.

O Quadro I apresenta as palavras com maior número de ocorrências, resultado da análise léxica realizada pela ferramenta *textalyser* nos resumos e nas suas palavras-chaves. Observa-se que o termo descritor “dados” e “*blockchain*” são os mais recorrentes nos trabalhos recuperados, incluindo os termos tecnologia, sistema, *bitcoin* e segurança. Interessante ressaltar que termos relacionados à proteção dos direitos autorais não emergem entre os 20 primeiros termos recorrentes.

Quadro I: Ocorrências de palavras nos resumos e nas palavras-chaves dos documentos

Resumo			palavras-chaves		
Palavra	Qtd	Freq	Palavra	Qtd	Freq
data	164	2.7%	data	79	4.5%
blockchain	160	2.6%	systems	52	3%
technology	77	1.3%	block	35	2%
system	56	0.9%	information	34	1.9%
distributed	39	0.6%	chain	34	1.9%
information	37	0.6%	management	33	1.9%
management	34	0.6%	distributed	31	1.8%
peer	34	0.6%	blockchain	29	1.7%
systems	32	0.5%	peer	27	1.5%
digital	31	0.5%	electronic	26	1.5%
bitcoin	30	0.5%	privacy	26	1.5%
solution	29	0.5%	security	25	1.4%
security	28	0.5%	digital	25	1.4%
privacy	28	0.5%	money	22	1.3%
transactions	26	0.4%	internet	19	1.1%
energy	26	0.4%	storage	19	1.1%
smart	25	0.4%	things	15	0.9%
research	24	0.4%	computing	14	0.8%
network	23	0.4%	computer	14	0.8%
financial	23	0.4%	networks	14	0.8%

Fonte: Dados da pesquisa.

Ao realizar mineração de texto na linguagem R para proporcionar a visualização dos dados mais frequentes nos resumos dos trabalhos recuperados, obteve-se a Figura I, que também corrobora com os dados demonstrados na análise anterior (QUADRO I).

de dados abordados nos documentos, na quarta e quinta coluna estão as técnicas ou os processos aplicados com a tecnologia *blockchain* e, o resumo dos processos apresentados pelos autores.

Observa-se que *blockchain* é aplicada em dados identificados pelos autores como: dados de pesquisa, dados digitais ou simplesmente dados, explicitando o contexto dos quais os dados estão inseridos (MA; HUANG; GAO, 2018; PĂNESCU; MANTA, 2018), músicas e vídeos (GAO; NOBUHARA, 2017; O'DAIR; BEAVEN, 2017), imagens de obras de arte (ZHAOFENG; WEIHUA; HONGMIN, 2018) e *timestamp* (GIPP et al.2017).

Os artigos descrevem e discutem técnicas e/ou processos utilizados com o *Blockchain* como a plataforma MyCelia para música, o CryptSubmit para manuscritos, contratos inteligentes e marcas d'água para obras de arte (APÊNDICE A).

Dentre os principais benefícios e oportunidades relatados destacam-se os relacionados à precisão e acessibilidade aos dados de direitos autorais, transparência nos processos, meios para evitar a falsificação de dados e o gerenciamento de direitos autorais.

5 CONSIDERAÇÕES FINAIS

A revisão sistemática de literatura na base de dados Scopus revelou que apesar de serem recuperados 80 trabalhos, apenas oito apresentam a aplicação de *blockchain* para proteção autoral de dados. Mesmo os trabalhos que discutem aspectos relacionados a proveniência de dados, esses descrevem a facilidade da integração de dados de diversas fontes, a entrega de informações pelas fontes, a origem dos dados, a verificação de adulteração dos dados identificando em quais dados foram feitas as transformações, no entanto, não vinculam com o direito autoral.

Conclui-se que o uso de *blockchain* no contexto de proteção de dados autorais não é foco da maioria das pesquisas recuperadas, situação que torna essa temática uma oportunidade para pesquisas em diversas áreas do conhecimento, visto que, ao utilizá-la pode-se contribuir na proteção e comprovação do direito autoral de dados.

REFERÊNCIAS

BARDIN, L. **Análise de conteúdo**. 4. ed. Lisboa: Edições 70, 1977.

CHEN, J.; XUE, Yi. Bootstrapping a blockchain based ecosystem for big data exchange. In: Big Data (BigData Congress), 2017 IEEE International Congress on. IEEE, 2017. p. 460-463.

GAO, Y.; NOBUHARA, H. A Decentralized Trusted Timestamping Based on Blockchains. **IEEJ Journal of Industry Applications**, Tokyo, v. 6, n. 4, p. 252-257, 2017.

GIPP, B. et al. CryptSubmit: introducing securely timestamped manuscript submission and peer review feedback using the Blockchain. In: Proceedings of the ACM/IEEE joint conference on digital libraries, 2017, Toronto. **Proceedings...** 2017. Disponível em: <<https://ieeexplore.ieee.org/document/7991588> >. Acesso em: 20 set. 2018.

- KARAFILOSKI, E.; MISHEV, A. Blockchain solutions for big data challenges: A literature review. In: Smart Technologies, IEEE EUROCON International Conference on IEEE, 17., 2017. **Procedin- gs...**2017. p. 763-768. Acesso em: 20 set. 2018.
- KIM, H. M.; LASKOWSKI, M. Toward an ontology-driven blockchain design for supply-chain provenance. **Intelligent System in Accounting Finance e Management**, Hoboken, v. 25, n. 2, p. 18-27, Jan./Mar. 2018. Disponível em: <<https://onlinelibrary.wiley.com/doi/full/10.1002/isaf.1424>>. Acesso em: 20 set. 2018.
- KIM, Y.; STANTON, J. M. Institutional and individual influences on scientists' data sharing practices. **Journal of Computational Science Education**, v. 3, p. 47-56, 2013. Disponível em: <https://www.researchgate.net/publication/239522383_Institutional_and_Individual_Influences_on_Scientists'_Data_Sharing_Practices>. Acesso em: 29 set. 2018.
- KIM, Y.; STANTON, J. M. Institutional and individual factors affecting scientists' data-sharing behaviors: A multilevel analysis. **Journal Of the Association for Information Science and Technology**, v. 67, n. 4, p. 776-799, 2016. Disponível em: <<https://onlinelibrary.wiley.com/doi/full/10.1002/asi.23424>>. Acesso em: 29 set. 2018.
- KITCHENHAM, B. **Procedures for performing systematic reviews**: joint technical report. Keele: Keele University, 2004.
- MA, Z.; HUANG, W.; GAO, H. Secure DRM Scheme Based on Blockchain with High Credibility. *Chinese Journal of Electronics*, v. 27, n. 5, p. 1025-1036, 2018.
- NOBLEJAS, C. J.; RODRÍGUEZ, A. P. Recuperación y visualización de información en Web of Science y Scopus: una aproximación práctica. **Investigación Bibliotecológica: archivonomía, bibliotecología e información**, México, v. 28, n. 64, p. 15-31, 2014. Acesso em: 20 set. 2018.
- O'DAIR M., BEAVEN Z. The networked record industry: How blockchain technology could transform the record industry. **Strategic Change**, Hoboken, v. 26, n. 5, p. 471-480, Sept. 2017. Disponível em: <<https://onlinelibrary.wiley.com/doi/full/10.1002/jsc.2147>>. Acesso em: 20 set. 2018.
- PĂNESCU, A-T.; MANTA, V. Smart Contracts for Research Data Rights Management over the Ethereum Blockchain Network. **Science & Technology Libraries**, p. 1-11, 2018.
- TENOPIR C. et al. (2015) Changes in data sharing and data reuse practices and perceptions among scientists worldwide. **PLoS ONE**, San Francisco, v. 10, n. 8, Aug. 2015. Disponível em: <<https://doi.org/10.1371/journal.pone.0134826>>. Acesso em: 16 fev. 2018.
- ZHAOFENG, M.; WEIHUA, H.; HONGMIN, G. A new blockchain-based trusted DRM scheme for built-in content protection. **EURASIP: Journal on Image and Video Processing**, v. 2018, n. 91, p. 1-12, Dec. 2018. Disponível em: <<https://link.springer.com/article/10.1186/s13640-018-0327-1>>. Acesso em: 20 set. 2018.

Apêndice A – Uso de blockchain para proteção autoral de dados

AUTORES	TÍTULO DO ARTIGO	TIPO DE DADO	TÉCNICA/PROCESSO	DESCRIÇÃO DO PROCESSO
O'DAIR; BEAVEN	The networked record industry: How blockchain technology could transform the record industry	Música	Plataforma Mycelia; banco de dados universal e único de direitos autorais de música; Aurovine; Global Repertoire Database (GRD); contratos inteligentes; micropagamentos; proof-of-work	Precisão e acessibilidade dos dados de direitos autorais; velocidade de pagamentos de royalties; e transparência da cadeia de valor.
GIPP et al.	CryptSubmit: Introducing Securely Timestamped Manuscript Submission and Peer Review Feedback Using the Blockchain	Timestamp	CryptSubmit - sistema de submissão de manuscritos para a publicação acadêmica	Sistema que cria automaticamente um timestamp descentralizado, à prova de falsificação e publicamente verificável para cada manuscrito submetido
PANESCU; MANTA	Smart Contracts for Research Data Rights Management over the Ethereum Blockchain Network	Dados de pesquisa	Solidity Contrato inteligente	Permite aos autores supervisionar como está sendo acessado e reutilizado.
MA; HUANG; GAO	Secure DRM scheme based on blockchain with high credibility	Dados digitais	Infraestrutura BaaS com plataforma DRM	Criação de blocos gênese, definição de mecanismos de consenso, inicialização e execução de nós, gerenciamento de wallet, gerenciamento de endereços, blockchain explorer
CHEN; XUE	Bootstrapping a Blockchain Based Ecosystem for Big Data Exchange	Dados	Ecosistema para troca de dados	Visa cultivar um ecossistema, dentro do qual todos os participantes podem cooperar para trocar dados de maneira peer-to-peer
GAO; NOBUHARA	A decentralized trusted timestamping based on blockchains	Dados digitais (documentos, música e vídeos)	Registro de data e hora	Usando blockchain expandindo o espaço de armazenamento de 40 bytes para N * 20 bytes.
ZHAOFENG; WEIHUA; HONGMIN	A new blockchain-based trusted DRM scheme for built-in content protection	Imagem de obras de arte	Marca d'água com esquema DRMChain	Esquema de gerenciamento de direitos digitais de imagens de obras de arte para detecção de uso incorreto, incorporamos informações corretas de obras de arte, como autor, RightHolder, data e local

Fonte: Dados da pesquisa.

pesquisa.