**ESTUDO DE CASO PARA MÉTODOS DE APLICAÇÃO DA LEI GERAL DE PROTEÇÃO DE DADOS**

Juan Cardoso da Silva, Ronaldo Celso Messias Correia

Departamento de Matemática e Computação (DMC)

Universidade Estadual Paulista “Júlio Mesquita Filho” (UNESP)

Faculdade de Ciências e Tecnologia (FCT)

Presidente Prudente - SP, Brasil

[**juan.c.silva@unesp.br**](mailto:juan.c.silva@unesp.br)**,** [**ronaldo.correia@unesp.br**](mailto:ronaldo.correia@unesp.br)

**Resumo – Com a ampliação das leis protetoras de dados digitais privados ao redor do mundo como G.D.P.R e a L.G.P.D. - Surgiu um espaço não explorado e sem atenção até as décadas recentes, a interseção das ciências humanas com as ciências exatas, a aplicação de Inteligências Artificiais para a manipulação desses dados dentro do contexto da área de direito para de alguma forma facilitar o cumprimento da lei em garantir a persistência dos dados e sua proteção. Neste trabalho será apresentado um processo de análise e identificação de dados privados por meio de diversos modelos de máquinas criados, com objetivo de analisar o potencial dessas máquinas ao identificar os dados como sua capacidade de recomendar encriptações como um método de proteção de tais dados.**

***Palavras Chaves –* Lei Geral de Proteção de dados (LGPD); Proteção de dados; Métodos de aplicação da LGPD, Manipulação e minimização de dados; Machine Learning; Inteligencia Artificial;**

# INTRODUÇÃO

Com a evolução da tecnologia focada em conectar computadores se tornou disponível para as pessoas nos anos 90, surgiu a World Wide Web e futuramente, nomeado Web 1.0 – Sendo um sucesso mundial e fornecendo diversos sites focados em encontrar informações espalhadas na internet, geralmente contendo sites estáticos e sem muitos detalhes e funcionalidades, também, essa versão da web fornecia atualizações para sistemas e softwares, tais como o rudimentar Windows Update do Windows 95 e sites de terceiros dedicados para jogos como Quakeworld (hoje em dia continuado por fãs, o web site original está fechado), nessa época a internet vivia sem rumo sobre sua funcionalidade concreta no futuro, até que em 1999 surgiu o primeiro uso do termo para sua segunda iteração, a Web 2.0

A Web 2.0 foi pensada na interatividade entre pessoas na internet além do acesso e busca de informações, com isso diversos sites novos começaram a nascer focados em fornecer serviços como os de construções de Wikis, Aplicações Web, Mashups, Redes Sociais, Eventos Colaborativos (caridades e doações para determinados motivos tais quais crowdsourcing e crowdfunding, ambos para financiar tipos de projetos, tais como Patreon e Kickstarter), Conteúdo gerado de usuários (Youtube), Blogs, Streaming Curadoria Social (Reddit e Instagram). Com esse serviço disponível agora em mãos dos usuários e popularizado ao ponto de dados sensíveis como cartões de créditos, emails e outras informações pessoais circulando, era inevitável a chegada de agentes indesejáveis atacando a rede em busca de acesso a esses dados, buscando lucro em tentativas de se apropriar desses dados.

Os ataques desses agentes eventualmente resultaram na criação de leis progenitoras, focadas em tratar ataques relacionados a privacidades, como a Lei de Combate a Crimes Cibernéticos (12.737/2012) criada para tratar o caso da Carolina Dieckmann. Em 2016 um evento no Reino Unido onde o Facebook indevidamente estava utilizando dados privados de alunos em uma universidade, acabou resultando na *General Data Protection Regulation* (GDPR), dois anos depois, o congresso brasileiro criou sua resposta em forma de lei, utilizando a lei 12.737/2012 e bases da GDPR para criar a Lei Geral de Proteção de dados (LGPD). Após sua implementação em maio de 2021, diversos métodos de implementação da lei foram adotados para respeita-la, de novas opções de gerenciamentos de cookies até como os dados são manipulados no em sistemas.

# FUNDAMENTAÇÃO TEÓRICA

AAA

# TRABALHOS RELACIONADOS

Nesta implementação da IBM [3], o QRadar funciona como um sistema automatizado na Cloud para detecção de irregularidades em uma base de dados, caso um cliente deseje seus dados sensíveis deletados, o QRadar vai realizar uma varredura para confirmar se os dados sensíveis foram realmente deletados e ativar um trigger caso este ainda esteja no sistema onde foi detectado. Nesta revisão de literatura, procura entender e ajudar a melhorar o conceito utilizado pela IBM permitindo não só essas detecções, como também avisando quando dados sensíveis não estão encriptados, realizar o descarte de dados sensíveis quando estes não forem descartados direito, invés de apenas avisar.

Neste trabalho de pesquisa [4], também da IBM, o modelo de ML utilizado por pelos pesquisadores usa escolhas binárias para os dados coletados serem minimizados de acordo com a GDPR, no caso, estão a utilizar dado de um hospital para escolher quais dados podem ser deletados baseado nos parâmetros escolhidos. O projeto a ser criado, propõe não só minimizar os dados como também encriptar os dados sensíveis, sendo uma melhora em relação a esta pesquisa da IBM.

# REFERÊNCIA BIBLIOGRÁFICA

**[1] Lei 'Carolina Dieckmann', que pune invasão de PCs, entra em vigor.** G1 Globo, São Paulo, dia 1, junho de 2013. Disponível em: <https://g1.globo.com/tecnologia/noticia/2013/04/lei-carolina-dieckmann-que-pune-invasao-de-pcs-passa-valer-amanha.html> Acesso em: 27, agosto de 2022

**[2]** Leal da Silva, Julia; **Tomada de Decisão Automatizada e Controle pela LGPD.** IAPD, 20, janeiro de 2021. Disponível em: <https://iapd.org.br/decisao-automatizada-lgpd-direito-aexplicacao/> Acesso em: 24, agosto de 2022.

**[3]** Fortunato, Caroline, **Using QRadar for LGDP**, 19, julho de 2019, Disponível em:<https://www.proof.com.br/wp-content/uploads/2019/08/Using-QRadar-for-LGPD.pdf>

**[4]** Goldsteen, Abigail; Ezov Gilad; Shmelkin, Ron; Moffie, Micha; Farkash, Ariel; **Data minimization for GDPR compliance in machine learning models**, p. 1-15, 2021, Disponível em: <<https://link.springer.com/article/10.1007/s43681-021-00095-8>>

**[5]** Roberto Fernandes Castilho, José; Andrade Gomes, Henrique; **Legislação Básica de Direito da Informática**, 2° edição reformulada e atualizada, São Paulo, Editora Pillares.

**[6] General Data Protection Regulation GDPR:** OJ L 119, 04.05.2016; cor. OJ L 127, 23.5.2018. Disponível em: <<https://gdpr-info.eu/>> Acesso em: 27, agosto de 2022

**[7]** Stallings, William; Brown, Lawrie; **Computer Security: Principles and Practice Second Edition**, 2° edição reformulada e atualizada, São Paulo, Pearson; 2ª edição.

**[8]** Mota, José; **Da Web 2.0 ao E-Learning 2.0: Aprender na rede,** p.1-199, 2009. Disponível em: <https://repositorioaberto.uab.pt/handle/10400.2/1381>

**[9]** Filipe Lima Rapôso, Cláudio; Melo de Lima, Haniel; Ferreira de Oliveira Junior, Waldecy; Aragão Ferreira Silva, Paola; Elaine de Souza Barros, Elaine; **LGPD - LEI GERAL DE PROTEÇÃO DE DADOS PESSOAIS EM TECNOLOGIA DA INFORMAÇÃO: Revisão Sistemática**, p. 1-10. 2019. Disponível em:<https://revistas.cesmac.edu.br/index.php/administracao/article/view/1035>

**[10]** Spadaccini de Teffé, Chiara; Viola, Mario; **Tratamento de dados pessoais na LGPD: estudo sobre as bases legais**, p. 1-38, 2020. Disponível em: <https://civilistica.emnuvens.com.br/redc/article/view/510>

**[11]** Tankard, Colin; **What the GDPR means for business**, p. 1-8, 2016. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1353485816300563?casa\_token=pjOHlq5iYsAAAAA:dhlGukGxSanwjnBY9aPtw36O1CslOJG1wZ7wTEryTJHi5QtFVM2G6kc8CwuemdVrUDRl2tRqJoA>

**[12]** Goldsteen, Abigail; Ezov Gilad; Shmelkin, Ron; Moffie, Micha; Farkash, Ariel; **Data minimization for GDPR compliance in machine learning models**, p. 1-15, 2021. Disponível em: <https://link.springer.com/article/10.1007/s43681-021-00095-8>

**[13]** Shanmugam, Divya; Shabanian, Samira; Diaz, Fernando; Finck, Michèle, Biega, Asia; **Learning to Limit Data Collectionvia Scaling laws: A computional Interpolation for the Legal Principle of Data Minimization**, p. 1-11, 2022. Disponível em: < https://arxiv.org/abs/2107.08096 >

**[14]** D. R. Ignatius Moses Setiadi, A. Faishal Najib, E. H. Rachmawanto, C. Atika Sari, K. Sarker and N. Rijati, "**A Comparative Study MD5 and SHA1 Algorithms to Encrypt REST API Authentication on Mobile-based Application**," 2019 International Conference on Information and Communications Technology (ICOIACT), 2019, pp. 206-211, doi: 10.1109/ICOIACT46704.2019.8938570. Disponível em: <https://ieeexplore.ieee.org/document/8938570>

**[15]** Boutaba, Raouf; Mohammad A., Salahuddin; Limam, Noura; Ayoubi, Sara; Shahriar, Nashid; Estrada-Solano Felipe; Caicedo M. Oscar; **A comprehensive survey on machine learning for networking: Evolution, applications and research opportunities**, p. 1-99, 2018. Disponível em:

<https://jisajournal.springeropen.com/articles/10.1186/s13174-018-0087-2#Sec2>

**[16]** Fatih, Ertam; Galip, Aydin; **Data classification with deep learning using TensorFlow**, p. 1-4, 2017. Disponível em:

<https://ieeexplore.ieee.org/document/8093521>

**[17]** Nguyen Quang-Hung; Hieu Doan; Nam Thoai; **Performance Evaluation of Distributed Training in TensorFlow 2**, p. 1-5, 2020. Disponível em: < https://ieeexplore.ieee.org/document/9353085 >

**[18]** PYTHON SOFTWARE FOUNDATION. **Python Language** Site: Documentation, 2022. Disponível em: <https://www.python.org/doc/> Acesso em: 24 de agosto de 2022.

**[19]** PANDAS **Pandas DataFrame** Site: About, 2022. Página sobre nós. Disponível em: <https://pandas.pydata.org/about/> Acesso em: 24 de agosto de 2022.

**[20]** PANDAS **Pandas DataFrame** Site: Documentation, 2022. Disponível em: <https://pandas.pydata.org/about/> Acesso em: 24 de agosto de 2022.

**[21]** PANDAS **10 minutes to pandas** Site: User Guide, 2022. Disponível em: <https://pandas.pydata.org/docs/user\_guide/10min.html#viewing-dataAcesso em: 24 de agosto de 2022.

**[22]** NUMPY PROJECT AND COMMUNITY. **What is Numpy?** Site: Documentation, 2022. Página de documentação. Disponível em:

<https://numpy.org/doc/stable/user/whatisnumpy.html> Acesso em: 24 de agosto de 2022.

**[23]** J. D. Hunter, "**Matplotlib: A 2D Graphics Environment**," in Computing in Science & Engineering, vol. 9, no. 3, pp. 90-95, May-June 2007, doi: 10.1109/MCSE.2007.55. Disponível em:<https://ieeexplore.ieee.org/document/4160265>

**[24]** MDN Web Docs, **Mozilla Developer Network Web Docs**, 2022. Disponível em: <https://developer.mozilla.org/pt-BR/>

**[25]** CLOUD GOOGLE. **Considerações sobre dados confidenciais em conjuntos de dados de aprendizado de máquina.** Site: Documentação, 2022. Disponível em: <<https://cloud.google.com/architecture/sensitive-data-and-ml-datasets>> Acesso em: 27, agosto de 2022.

**[26]** D Rachmawati, J T Tarigan and A B C Ginting; **A comparative study of Message Digest 5(MD5) and SHA256 algorithm**, p. 1-7, 2018. Disponível em <<https://iopscience.iop.org/article/10.1088/1742-6596/978/1/012116/meta>>

**[27]** Song, Congzheng; Ristenpart, Thomas; Shmatikov, Vitaly; **Machine Learning Models that Remember Too Much**, p. 1-15, 22, setembro de 2017. Disponível em: <<https://arxiv.org/pdf/1709.07886.pdf>>

**[28]** Vieira Souza, Iuri; **Aplicações de software desenvolvidas no contexto da  
Inteligência Artificial (IA), Machine Learning e Big  
Data e o direito dos cidadãos de acordo com a Lei  
Geral de Proteção de Dados (LGPD)**, p. 1-83, 2021. Disponível em: < https://bdm.unb.br/bitstream/10483/30275/1/2021\_IuriSousaVieira\_tcc.pdf>

**[29]** Goldsteen, Abigail; Ezov Gilad; Shmelkin, Ron; Moffie, Micha; Farkash, Ariel; **Data minimization for GDPR compliance in machine learning models**, p. 1-15, 2021, Disponível em: <<https://link.springer.com/article/10.1007/s43681-021-00095-8>>

[30] K. Pavani and P. Sriramya, "Enhancing Public Key Cryptography using RSA, RSA-CRT and N-Prime RSA with Multiple Keys," 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), 2021, pp. 1-6, doi: 10.1109/ICICV50876

.2021.9388621. Disponível em: <https://ieeexplore.ieee.org/document/9388621>