
O ALINHAMENTO ENTRE LEARNING ANALYTICS E A GENERAL DATA PROTECTION REGULATION: UMA REVISÃO SISTEMÁTICA DE LITERATURA

THE ALIGNMENT BETWEEN LEARNING ANALYTICS AND THE GENERAL DATA PROTECTION REGULATION: A SYSTEMATIC LITERATURE REVIEW

LA ALINEACIÓN ENTRE LEARNING ANALYTICS Y EL REGLAMENTO GENERAL DATA PROTECTION REGULATION:
UNA REVISIÓN SISTEMÁTICA DE LA LITERATURE

Rafael Vescovi Bassani¹, Silvio César Cazella²

RESUMO

O rápido crescimento da modalidade de educação a distância propiciou que pesquisadores apresentassem estudos variados relacionados ao tema. Junto a essas pesquisas, emergem conceitos, como o de Learning Analytics (LA), uma área que se propõe a medir, coletar, analisar e relatar dados de discentes em seus contextos, o que traz à tona nova preocupação, relacionada à proteção, à privacidade e ao correto uso desses dados. A União Europeia já se encontra legislando sobre a proteção de dados pessoais, com um amplo Regulamento Geral sobre a Proteção de Dados (GDPR). Nesse sentido, este artigo tem por objetivo apresentar o resultado de uma Revisão Sistemática da Literatura (RSL), que buscou identificar pesquisas acadêmicas relacionadas à temática Learning Analytics (LA) e General Data Protection Regulation (GDPR). Após a aplicação dos critérios de inclusão e exclusão sobre os artigos obtidos, foram selecionados dez textos, cuja análise possibilitou concluir que há um alinhamento entre os conceitos de LA e GDPR e que o LA deve seguir as orientações do GDPR.

PALAVRAS-CHAVE: Revisão sistemática de literatura, Learning Analytics e General Data Protection Regulation.

ABSTRACT

The growth of the distance education modality allows researchers to present varied studies related to the theme. Together with these studies emerge concepts such as Learning Analytics (LA). The LA is an area that will analyze, measure, collect and relate student data in their contexts. However, the use of this data brings with it a new concern regarding the protection, privacy and correct use of the data. The European Union already finds personal data protection legislation with a broad General Data Protection Regulation (GDPR). This article aims to present the result of a Systematic Literature Review that search to identify academic research related to Learning Analytics (LA) and General Data Protection Regulation (GDPR). After applying the inclusion and exclusion requirements of the selected articles ten articles were selected for analysis. With the analysis it is possible to identify an alignment between the LA and GDPR showing that LA should follow the guidelines of GDPR.

KEYWORDS: Systematic literature review, Learning Analytics and General Data Protection Regulation.

¹ Doutorando do Programa de Pós-graduação em Informática em Educação (PPGIE UFRGS). Professor na Universidade do Vale do Rio dos Sinos (UNISINOS). São Leopoldo, RS - Brasil. E-mail: rvbassani@gmail.com

² Doutor em Ciência da Computação (UFRGS). Professor da (UFRGS), (UFCSPA) e do Programa de Pós-graduação em Informática em Educação (PPGIE UFRGS). Porto Alegre, RS - Brasil. E-mail: silvioc@ufcspa.edu.br
Submetido em: 22/03/2020 - Aceito em: 04/02/2021

RESUMEN

El rápido crecimiento de la modalidad de educación a distancia ha permitido a los investigadores presentar diversos estudios relacionados con el tema. Junto con estos estudios, surgen conceptos como Learning Analytics (LA), que es un área que tiene como objetivo medir, recopilar, analizar e informar datos de los estudiantes en sus contextos. Sin embargo, el uso de estos datos de los estudiantes plantea una nueva preocupación relacionada con la protección, la privacidad y el uso correcto de los datos. La Unión Europea ya está legislando sobre la protección de datos personales con un amplio General Data Protection Regulation (GDPR). Este artículo tiene como objetivo presentar el resultado de una Revisión sistemática de literatura (RSL) que buscó identificar investigaciones académicas relacionadas con Learning Analytics (LA) y el General Data Protection Regulation (GDPR). Después de aplicar los criterios de inclusión y exclusión en los artículos obtenidos, se seleccionaron diez artículos para su análisis. Con el análisis fue posible concluir que existe una alineación entre los conceptos de LA y GDPR y que LA debe seguir las pautas de GDPR.

PALAVRAS-CLAVE: Revisión sistemática de la literatura, Learning Analytics y General Data Protection Regulation.

1 INTRODUÇÃO

A educação a distância, para Hack (2011), é uma modalidade de ensino que permite atingir um número significativo de pessoas e rompe com a forma tradicional de ensino. Com o crescimento da Educação a Distância, é observado um aumento no volume de dados disponíveis (armazenados) em bancos de dados. Segundo Castilhos (2009), existe uma produção massiva dos dados que compõem as informações, os quais crescem cerca de 30% ao ano, de forma exponencial, e possibilitam a realização de pesquisas nessa área (CHATTI et al., 2012). De acordo com Koedinger et al. (2008), todos esses dados podem ser analisados por educadores, pesquisadores e gestores.

Nesse cenário, surgem conceitos, como o de Learning Analytics (LA), que, conforme Siemens e Long (2011), é a medição, coleta, análise e interpretação dos dados em ambientes educacionais. Chatti et al. (2012) esclarecem que diferentes definições foram propostas para o termo LA, mas todas possuem ênfase na conversão de dados educacionais em informações que irão permitir a realização de ações para apoiar melhorias no processo de aprendizagem. Para Johnson et al. (2011), o LA refere-se à interpretação de um grande volume de dados produzidos por alunos, com o objetivo de avaliar o progresso acadêmico. Moissa et al. (2015), por sua vez, trazem que o LA vai buscar entender o processo de ensino-aprendizagem como um todo ou desenvolver ferramentas para auxiliar nesse processo, em uma abordagem mais voltada para os fatores humanos e utilizando técnicas de Educational Data Mining (EDM) como um instrumento. Pesquisadores que participaram da primeira edição da Conferência Internacional em Análise de Aprendizado e Conhecimento (International Conference on Learning Analytics and Knowledge – LAK)³ destacam que a área de Learning Analytics irá

³ 1st International Conference on Learning Analytics and Knowledge - Disponível em: <https://tekri.athabascau.ca/analytics/>

medir, coletar, analisar e relatar dados de estudantes e seus contextos. Dessa forma, o LA visa tanto compreender o aprendizado do aluno como melhorar o ambiente em que este está inserido (2014). Siemens e Baker (2012) ainda indicam o uso de LA para proporcionar experiências personalizadas e alinhar estratégias pedagógicas de forma assertiva para os diferentes perfis de alunos.

Coletar e trabalhar com os dados dos alunos, entretanto, pode não ser algo tão simples ou totalmente disponível. Alguns países, preocupados com a privacidade e proteção de dados pessoais de seus cidadãos, aprovaram algum tipo de regulação. Para Altman (1975), a privacidade é um fenômeno culturalmente difundido, ou seja, pessoas em diferentes culturas são universalmente conscientes e capazes de regular a privacidade interpessoal, mas seus comportamentos específicos variam de cultura para cultura.

Observando esse cenário, a União Europeia está, desde 1995, atuando e legislando sobre a proteção de dados pessoais, tendo aprovado, em 2016, um amplo Regulamento Geral sobre a Proteção de Dados⁴. Trata-se de um regulamento do direito europeu, cuja implementação iniciou em 25 de maio de 2018, que legisla sobre a privacidade e proteção de dados pessoais, aplicável a todos os indivíduos na União Europeia e Espaço Econômico Europeu.

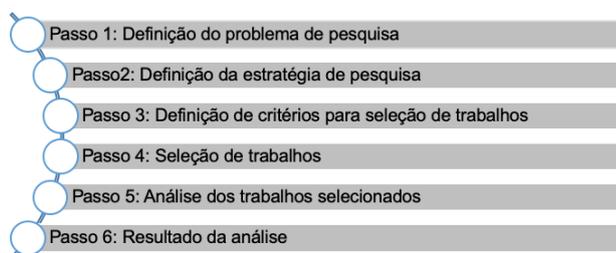
Nesse contexto, este artigo apresenta o resultado de uma Revisão Sistemática de Literatura (RSL), em que se buscou identificar pesquisas relacionadas com a temática de Learning Analytics (LA) e aplicação do General Data Protection Regulation (GDPR). A correlação entre os temas de LA e GDPR justifica-se pelo fato de a GDPR nortear as questões sobre a privacidade dos dados pessoais e a LA estudar os sistemas e as práticas de análise de dados. Para Thorpe et al. (2005), a Revisão Sistemática de Literatura (RSL) caracteriza-se por empregar uma metodologia de pesquisa com rigor científico e de grande transparência, cujo objetivo visa minimizar o enviesamento da literatura, na medida em que é feita uma pesquisa exaustiva dos textos publicados sobre o tema em questão. Com relação à estrutura deste artigo, apresenta-se em quatro seções: (1) Introdução; (2) Método; (3) Resultados; (4) Considerações Finais e, por fim, Referências.

⁴ Regulamento Geral sobre a Proteção de Dados N° 679/2016 - Disponível em: <https://eur-lex.europa.eu/eli/reg/2016/679/oj?locale=pt>

2 MÉTODO

O presente estudo caracteriza-se como uma Revisão Sistemática da Literatura (RSL), que, segundo Kitchenham (2007), emprega um processo para identificar, avaliar e interpretar evidências científicas relacionadas ao tema específico pesquisado. A esse conceito, somam-se Dresch et al. (2015), para quem a RSL corrobora o mapeamento e possibilita encontrar, avaliar e, também, identificar lacunas a serem preenchidas e, assim, proporciona uma visão robusta acerca de determinada temática, possibilitando aos pesquisadores estar a par de estudos atuais. O artigo, destaca-se, segue o processo de execução de uma Revisão Sistemática de Literatura (RSL) empregado por Soni e Kodali (2011), com (06) seis passos, descritos na Figura 1, que segue.

Figura 1 – Passos de uma RSL

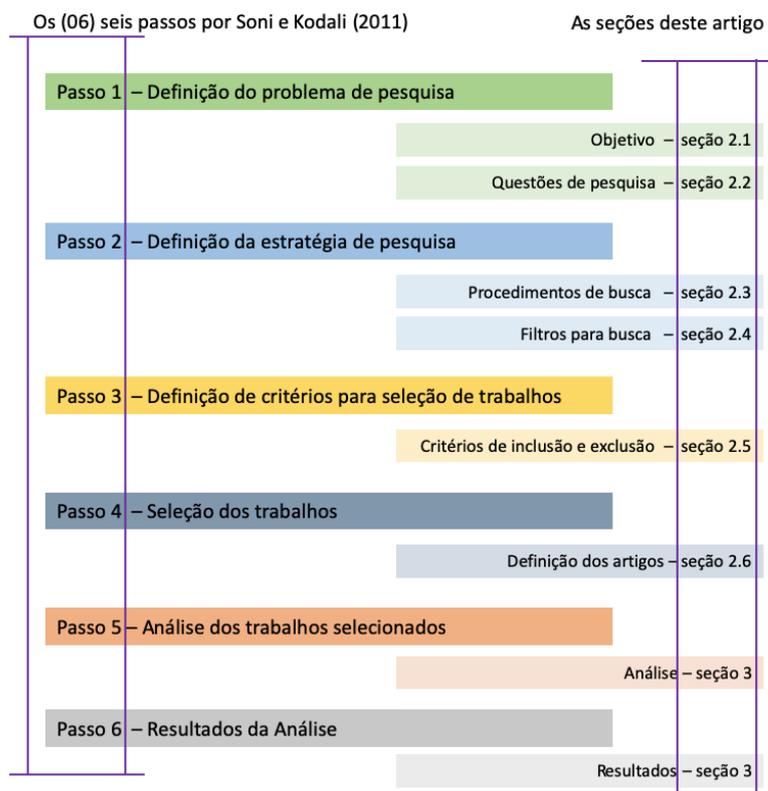


Fonte: Adaptado de Soni e Kodali (2011)

Como primeiro passo, a “definição do problema de pesquisa”, foi identificado o objetivo central da pesquisa, que busca identificar pesquisas relacionadas com a temática de Learning Analytics (LA) e aplicação do General Data Protection Regulation (GDPR), assim como buscou-se o alinhamento entre LA e a GDPR, além da definição das (04) quatro questões de pesquisa que norteiam esta investigação. No passo 2, a “definição da estratégia de pesquisa”, elucidaram-se o procedimento de busca dos artigos científicos e os filtros realizados nessa busca. O passo 3 consistiu na “definição de critérios para seleção de trabalhos”, com a demonstração dos critérios de inclusão e exclusão dos artigos científicos, selecionados e apresentados no passo 4, a “seleção de trabalhos”. No passo 5, realizou-se a “análise dos trabalhos selecionados” e, por fim, no 6, denominado “resultados da análise”, procedeu-se ao fechamento da pesquisa.

Para apoio na elaboração desta RSL, foi construído um protocolo com: (a) objetivo; (b) questões de pesquisa; (c) procedimento de busca; (d) filtros para busca e, por fim, (e) critérios de inclusão e critérios de exclusão. Ainda, para corroborar o alinhamento entre o processo de execução da RSL proposto por Soni e Kodali (2011) e esta pesquisa, apresenta-se a Figura 2.

Figura 2 - Alinhamento entre os passos da RSL de Soni e Kodali (2011) e seções deste artigo



Fonte: Elaborado pelo autor

No lado esquerdo da Figura, estão os (06) seis passos para execução da RSL citados pelos dois autores. No lado direito, dispõem-se as seções desta pesquisa e os pontos em que elas se alinham aos passos propostos por Soni e Kodali (2011).

2.1 Objetivo

Como primeira definição, a escolha do objetivo da pesquisa - identificar pesquisas relacionadas com a temática de Learning Analytics (LA) - e o alinhamento entre LA e a GDPR.

2.2 Questões da Pesquisa

A Tabela 1, que segue, apresenta as (04) quatro questões de pesquisa (QP) definidas para conduzir esta Revisão Sistemática de Literatura.

Tabela 1: Questões de pesquisa

QUESTÃO DE PESQUISA	PERGUNTAS
QP1	Quais são os principais conceitos que estão sendo relacionados a Learning Analytics nestas pesquisas?
QP2	Quais são as tecnologias de informação e comunicação (TICs) que estão sendo aplicadas à temática de Learning Analytics?
QP3	Como o conceito de GDPR está sendo aplicado nestas pesquisas?
QP4	Como a aplicação da GDPR está contribuindo para a gestão universitária?

Fonte: Elaborado pelo autor

2.3 Procedimento de busca de artigos científicos

Para esta seção, levou-se em consideração a busca por artigos acadêmicos com revisão por pares, nos portais SCIENCE DIRECT⁵, WEB OF SCIENCE⁶ E SCOPUS⁷, publicados entre os anos de 2015 e 2019. A *string* de busca definida foi: (Learning Analytics or Academic Analytics) and (General Data Protection Regulation or GDPR). Após essa definição, foi realizada a recuperação de artigos nos respectivos portais. Foram recuperados, ao todo, (42) quarenta e dois artigos, sendo (26) vinte e seis pelo portal SCIENCE DIRECT, (08) oito pelo portal SCOPUS e, por fim, (08) oito pelo portal WEB OF SCIENCE.

Após a recuperação dos artigos, eles foram exportados e salvos em extensão BIBTEX ou CSV. A seguir, foram lançados na plataforma RAYYAN QCRI⁸, que é 100% gratuita na web e apresentada com a finalidade de ajudar os autores nas revisões sistemáticas. A escolha dessa plataforma ocorreu devido à facilidade na sua utilização.

2.4 Filtros para busca

Os filtros para busca aplicados foram, inicialmente, de período - nesse caso, aplicados sobre os anos de “2015 até 2019”; o idioma identificado foi o da “língua Inglesa”; o tipo de publicação selecionado foi o de “somente artigos” e o tipo de acesso, “*open access*”. Os filtros para busca são apresentados na Tabela 2, que segue.

⁵ Acesso disponível: <https://www.sciencedirect.com>

⁶ Acesso disponível: <http://www.webofknowledge.com>

⁷ Acesso disponível: <https://www2.scopus.com/standard/marketing.uri>

⁸ Acesso disponível: <https://rayyan.qcri.org/welcome>

Tabela 2: Filtros para busca

FILTROS PARA BUSCA	INFORMAÇÕES
PERÍODO	(2015, 2016, 2017, 2018 e 2019).
IDIOMA	Em inglês
TIPOS DE PUBLICAÇÃO	Somente artigos
TIPO DE ACESSO	<i>Open access</i>

Fonte: Elaborado pelo autor

Destaca-se que o período de 5 anos (2015-2019) foi aplicado nos filtros de busca no intuito de se ter acesso a publicações mais recentes, assim como o idioma, inglês, por alcançar um número maior de publicações sobre o assunto abordado e o tipo de acesso, *open access*, pela acessibilidade aos artigos.

2.5 Critérios de inclusão e exclusão

Todos os critérios de inclusão e exclusão descritos na Tabela 3, que segue, foram aplicados na seleção dos artigos identificados. Inicialmente, foram analisados apenas os resumos, e, quando da ocorrência de dúvida se deveria ser incluído ou excluído, a consulta ao texto completo era realizada. Caso algum dos artigos se encaixasse diretamente em um dos 4 critérios de exclusão, era automaticamente recusado.

Tabela 3: Critérios de inclusão e exclusão

Nº	CRITÉRIOS DE INCLUSÃO	Nº	CRITÉRIOS DE EXCLUSÃO
1	Artigos publicados em periódicos	1	Artigos em duplicação
2	Artigos publicados no período entre janeiro de 2015 e dezembro de 2019	2	Artigos que apresentam revisões sistemáticas e mapeamentos sistemáticos
3	Artigos publicados no idioma inglês	3	Não se trata de um artigo, embora esteja classificado como tal em um periódico (editoriais, ensaio, <i>reviews</i> de livros...)
4	Periódicos revisados por partes	4	Artigos sem o foco da pesquisa

Fonte: Elaborado pelo autor

2.6 Seleção dos Artigos

Após a realização do procedimento de busca, da seleção para exportação dos artigos, da aplicação dos filtros para busca e da aplicação dos critérios de inclusão e exclusão, realizou-se a definição dos artigos a serem analisados. Nessa etapa, foi realizada a leitura completa de cada um dos artigos, aprimorando o foco desta pesquisa. Dessa análise, (32) trinta e dois textos foram excluídos e (10) dez, selecionados. A Figura 3, que segue, apresenta a síntese do processo de seleção dos artigos.

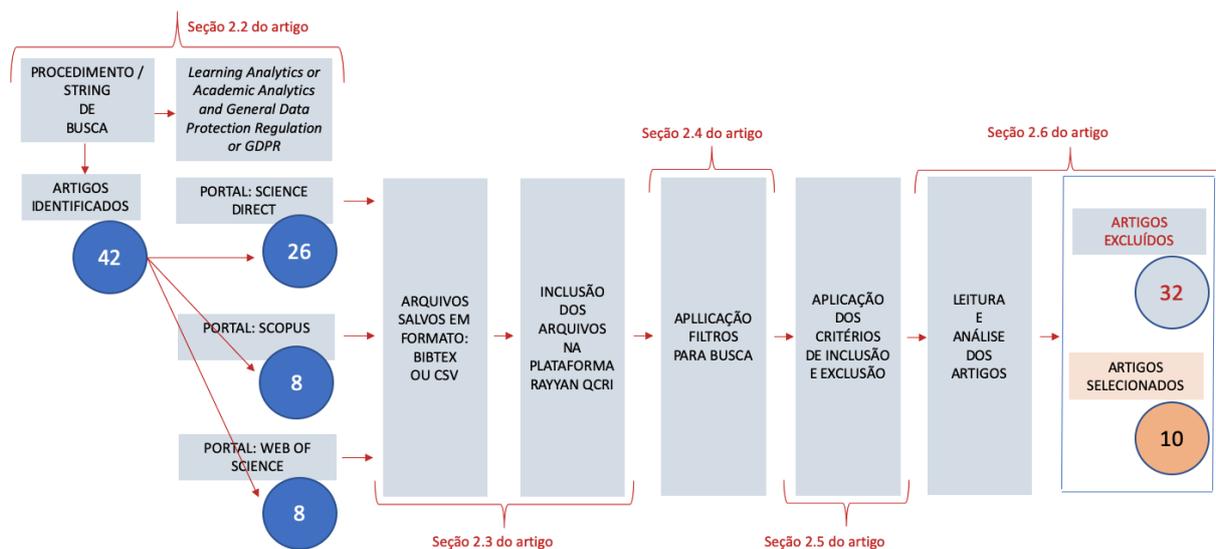


Figura 3 – Síntese do processo de seleção dos artigos

Fonte: Elaborado pelo autor

Os artigos selecionados encontram-se detalhados na Tabela 4, a seguir, em que são apresentados os títulos dos artigos, seus respectivos autores e os anos das publicações. As indicações dos artigos respondendo às QP estão apresentadas por cores: em azul, as indicações dos artigos que respondem à QP1; em cinza, os artigos que respondem à QP2; em verde, aqueles que respondem à QP3; e, em vermelho, os que respondem à QP4. As indicações das bases de fonte de busca dos artigos selecionados também são apresentadas por cores: em amarelo, as indicações dos artigos selecionados pelo portal WEB OF SCIENCE; em marrom, pelo portal SCOPUS; e, em rosa, pelo portal SCIENCE DIRECT. Desataca-se que há artigos que foram selecionados pelos (03) três portais.

Tabela 4: Artigos selecionados

Nr.	Artigo(s) / Autor(res)	Ano Publicação:	QP1	QP2	QP3	QP4	Web of Science	Scopus	Science Direct
1	Practical ethics for building learning analytics Autores: KITTO, K.; KNIGHT S.	2019	x	x	x	x	x	x	x
2	The ICO and Artificial Intelligence: The role of Fairness in the GDPR Framework Autor: BUTTERWORTH, M.	2018		x	x		x	x	
3	Fairness and Transparency of Machine Learning for Trustworthy Cloud Services Autores: ANTUNES, N.; et al.	2018		x	x			x	
4	Privacy Regulation Aware Process Mapping in GEO – Distributed Cloud Data Centers Autores: CHI ZHOU, A.; et al.	2019		x	x			x	
5	The Influence of Data Protection and Privacy Frameworks on the Design of Learning Analytics Systems Autores: HOEL, T.; GRIFFITHS D.; CHEN, W.	2017	x		x	x		x	
6	Data analytics in a privacy-concerned world Autores: WIERINGA, J.; et al.	2019			x				x
7	The Technology Foresight Activities of European Union Data Protection Authorities Autor: WILLS, D. B.	2016		x					x
8	Why Research May no Longer be the Same: About the Territorial Scope of the New Data Protection Regulation Autor: KINDT, E. J.	2016		x					x
9	Normative Challenges of Identification in The Internet of Things: Privacy, Profiling, Discrimination, and the GDPR Autor: WACHTER, S.	2018		x	x				x
10	Dualism in Data Protection: Balancing the Right to Personal Data and the Data Property Right Autores: YU, X.; ZHAO, Y.	2019			x				x

Fonte: Elaborado pelo autor

3 RESULTADOS

Finalizada a apresentação da introdução e dos materiais e métodos, pode-se passar para a discussão dos resultados.

De posse das informações coletadas e analisando a Tabela 4, percebe-se que, dentre as (04) quatro QP (questões de pesquisa), a QP1 é respondida por (03) três artigos, a QP2, por (07) sete artigos, a QP3, por (08) oito artigos e, finalizando, a QP4, por (02) dois artigos. O ano de 2019 é o que apresenta o maior número de artigos selecionados - ao todo são (04) quatro artigos -, 2018 apresenta (03) três artigos selecionados, 2017, (01) um artigo e, por fim, o ano de 2016, (02) dois artigos selecionados. O portal SCIENCE DIRECT é o que apresentou maior número de artigos selecionados - foram (06) seis ao todo.

Com relação à QP1, que se caracterizou por identificar os principais conceitos relacionados a Learning Analytics, foram identificados (02) dois artigos que sustentam essa resposta.

Para Kitto e Knight (2019), o Learning Analytics, desde a sua criação, está diretamente relacionado ao conceito de ética. Os autores destacam o surgimento de inúmeras listas de verificação e estruturas, sustentando que a privacidade do aluno seja respeitada e que possíveis danos sejam evitados. Contudo, também evidenciam a ocorrência de tensões e lacunas nas estruturas éticas e nas listas de verificação existentes para apoiar o desenvolvimento ético e a implementação da análise de aprendizado. Por fim, finalizam seus estudos afirmando ser provável que uma multiplicidade de abordagens éticas seja útil para entender como desenvolver e implementar análises de aprendizado de forma ética em contextos práticos.

Hoel, Griffiths e Chen (2017) consideram que o Learning Analytics fica direcionado ao conceito e ao objetivo da implantação do GDPR EU. O GDPR abre um cenário complexo de questões de privacidade e políticas que, por sua vez, influenciam como os sistemas e práticas de análise de aprendizado são projetados. Esses três autores afirmam, ainda, que, ao desenvolver uma lista mais detalhada de requisitos de projeto para sistemas de LA, o ponto de partida deve estar definido pela estrutura de privacidade usada pela legislação vigente.

Nos estudos, ficaram claros alguns pontos a serem apresentados nesta análise. Inicialmente, a relevância do tema Learning Analytics e, posteriormente, a aproximação da construção do LA com questões éticas e a observação de que requisitos legais são indispensáveis para sua aplicação e divulgação dos resultados.

A QP2 trouxe como questão quais são as tecnologias de informação e comunicação (TICs) que estão sendo aplicadas à temática de Learning Analytics. Nesse quesito, foram selecionados (07) sete artigos.

Kitto e Knight (2019) sustentam que a Inteligência Artificial (IA) e a Análise de Dados (AIDA) estão cada vez mais entrando na educação, com a inclusão de testes padronizados. Nesse sentido, os autores propõem um banco de dados piloto aberto que liste os casos enfrentados pelos construtores de sistemas de LA como um método para orientar os especialistas em ética. Isso proporcionaria um espaço intermediário no qual os construtores técnicos de sistemas poderiam interagir mais profundamente com os envolvidos com políticas, leis e ética para, assim, trabalhar de forma adequada na construção de LA.

A IA é, ainda, citada em outro artigo. Butterworth (2018) apresenta o Information Commissioner's Office (ICO). O ICO estuda as implicações da inteligência artificial para a proteção de dados pessoais e considera o uso de algoritmos de aprendizado de máquina para analisar um conjunto de dados e extrair correlações como uma reversão da análise de dados tradicional, em que a hipótese seria escolhida primeiro e, depois, os dados consultados para testar a hipótese.

Seguindo na apresentação das TIC(s), o projeto ATMOSPHERE é trazido por Antunes et al. (2018). Esse projeto tem como objetivo o desenvolvimento de um ecossistema para apoiar o *design* e o desenvolvimento de serviços em nuvem confiáveis da próxima geração, além de um conjunto de recursos federados e híbridos intercontinentais. Considera, também, a confiabilidade dependente de muitas propriedades, como segurança, confiabilidade e garantia de privacidade.

Chi Zhou et al. (2019) apresentam o Amazon EC2 e o Windows Azure para aplicação de método de otimização em duas nuvens públicas populares. O Amazon EC2 é um serviço via web que disponibiliza capacidade computacional redimensionável na nuvem, enquanto o Windows Azure é uma plataforma destinada à execução de aplicativos e serviços, baseada nos conceitos da computação em nuvem. Segundo os autores, como o desempenho da rede é um fator-chave para o desempenho do mapeamento de processos, foram estudados, empiricamente, os recursos de desempenho de rede dos datacenters no ambiente em nuvem distribuído geograficamente. Nesse estudo, os pesquisadores afirmam que os datacenters continuarão a ser distribuídos geograficamente à medida que volumes crescentes de dados com diferentes requisitos de regulamentação de privacidade estão sendo gerados e armazenados.

O tema das plataformas de pesquisa aparece também em outro artigo selecionado. Kindt (2016) traz que, em uma sociedade da informação digital inovadora, a pesquisa desempenha um papel fundamental. A coleta de dados pode ser feita a partir de sensores, sites públicos ou em colaboração com outras partes, organizadas, em particular, nas plataformas e em torno delas. Nesse sentido, o autor esclarece que as plataformas de pesquisa são um dos meios utilizados pelos pesquisadores para alcançar um grupo maior de usuários ou estimular a pesquisa ou permitir a comparação, por exemplo, de resultados de testes de algoritmos. Dentre as TIC(s), Kindt (2016) faz relação à utilização das plataformas BEAT PLATFORM; IOT LAB e Fed4Fire.

A Internet of Things (IOT) aparece em dois artigos identificados. O primeiro é através de Wills (2016), que comenta sobre as tecnologias relacionadas a cidades inteligentes, carros e residências envolvendo a proliferação de sensores e atuadores em todo o ambiente e a interconexão desses dispositivos com cada um e com o ambiente online. O autor ainda traz que a IOT é um ponto de intenso contato entre a tecnologia da informação em rede e o espaço físico. A conclusão do artigo é de que a tecnologia é uma área em que existe um alto nível de variação entre recursos e experiência.

O segundo artigo selecionado é de Wachter (2018). De acordo com o autor, a IOT é definida por conexões e serviços vinculados, adaptados aos requisitos específicos dos usuários. Os dispositivos de IOT coletam grandes quantidades de dados e oferecem maior percepção da vida dos usuários. A IOT exige identidades únicas de usuário e dispositivo para fornecer bancos de dados e serviços vinculados e personalizados para os usuários. Existe, nesse cenário, uma tensão fundamental entre duas forças nos sistemas de IOT: a privacidade do usuário e a tecnologia de identificação usada para vincular e personalizar serviços. Dessa forma, a pesquisa indica o surgimento de vários sistemas de autenticação (OpenID, PKI, SAML e Ucode) que suportam esse emparelhamento de gerenciamento de identidades e controle de acesso.

Da leitura dos artigos, ficou evidenciado que as tecnologias de informação e comunicação (TICs) aplicadas à temática de Learning Analytics passam pela aplicação de IA somadas a uma preocupação junto aos construtores de LA quanto à ética e à proteção de dados pessoais. O *design* do LA deve seguir padrões definidos e garantir segurança e privacidade dos dados. Quanto às pesquisas realizadas em atividades na nuvem, são apresentados o projeto ATMOSPHERE (ecossistema para apoiar o *design* e o desenvolvimento de serviços em nuvem) e o Amazon EC2 (serviço via web que disponibiliza capacidade computacional na nuvem). Quanto às plataformas, inicialmente, a destinada à execução de aplicativos e serviços é a Windows Azure, baseada nos conceitos da computação em nuvem. Posteriormente, as plataformas de pesquisa, em que são indicadas a utilização das BEAT PLATFORM; IOT LAB e Fed4Fire. Por fim, das TICs aplicadas à temática de Learning Analytics, aparece a IOT, interconexão desses dispositivos com cada um e com o ambiente online. Ela exige serviços personalizados e, ao mesmo tempo, o cuidado junto à privacidade do usuário, além da indicação de sistemas de autenticação, como (OpenID, PKI, SAML e Ucode).

A QP3, que pretendia averiguar como o conceito de GDPR estava sendo aplicado nessas pesquisas, possibilitou a identificação de (08) oito artigos que sustentam as respostas, apresentadas na sequência.

Para Kitto e Knight (2019), o GDPR adota uma definição ampla de dados pessoais, considerando-o como qualquer informação que permita a identificação de um indivíduo. Como exemplos, os autores citam nome, endereço, fotos, endereços IP, e-mails, postagens em sites de redes sociais e, até mesmo, informação médica. Os pesquisadores resgatam uma parte da legislação com maior destaque, identificando o papel dos controladores de dados (que coletam os dados de residentes da UE) e dos processadores de dados (que processam os dados coletados pelos controladores). O GDPR fornece um conjunto de regras que devem ser seguidas para proteger os direitos dos titulares de dados que geram dados pessoais através de suas ações no mundo. Essa estrutura necessitará de proteção de dados por projeto, exigindo que os controladores de dados pessoais implementem medidas técnicas e organizacionais apropriadas para garantir que apenas sejam processados aqueles dados pessoais necessários para cada finalidade específica.

Os requisitos aprimorados de responsabilidade e transparência do GDPR são citados por Butterworth (2018). Conforme o autor, os requisitos do GDPR representam desafios técnicos para desenvolvedores de IA e ajudam a mitigar alguns efeitos potencialmente adversos da IA em um contexto de dados pessoais. O artigo apresenta a ICO, que estuda as implicações da IA para a proteção de dados, e esclarece como o processamento de dados pessoais usando IA se encaixa na estrutura do GDPR.

O projeto ATMOSPHERE é apresentado por Antunes et al. (2018) e possui como objetivo desenvolver um ecossistema para apoiar o *design* e o desenvolvimento de serviços em nuvem. Nesse sentido, os autores trazem que o (GDPR) exige que as organizações tomem as medidas apropriadas para proteger os dados dos indivíduos e os usem de maneira a preservar a privacidade, justa e transparente.

O GDPR, de acordo com Chi Zhou et al. (2019), protege qualquer informação pessoal de europeus que possa ser usada para identificar, direta ou indiretamente, um indivíduo. Isso significa que, mesmo que os dados sejam anonimizados/criptografados, mesmo que a identidade pessoal possa ser revelada a partir deles, eles ainda estarão sob a proteção da legislação (GDPR EU).

Segundo Hoel, Griffiths e Chen (2017), à medida que a análise da aprendizagem nas escolas, universidades e local de trabalho começa a aumentar, as preocupações com a privacidade e a proteção de dados também aumentam. Nesse sentido, os autores afirmam que, dentre as legislações vigentes, o GDPR EU é o mais atualizado em relação aos desafios do mundo digital (regras relativas à notificação de violações, tomada e definição de perfil automatizadas e portabilidade de dados, como exemplos). O artigo faz menção a outra publicação de Hoel e Chen (2016), em que os desenvolvedores de LA utilizaram o modelo de

ciclo de vida do processo de LA da organização internacional de padronização ISO / IEC JTC1 / SC36 como modelo para discutir como os requisitos do GDPR influenciariam o desenvolvimento de sistemas. Hoel, Griffiths e Chen (2017) citam que o GDPR está sozinho na promoção do princípio de proteção de dados por *design* e por padrão. E, por fim, os autores esclarecem que o primeiro passo na cadeia de etapas do processamento de dados pessoais é a responsabilidade da coleta, pois oferece oportunidades únicas para a

implementação de medidas de privacidade no início do processo, o que pode garantir “privacidade por *design*”, conforme o GDPR, artigo 25.

O artigo 4 da legislação vigente do GDPR é resgatado por Wieringa et al. (2019). Dados pessoais são definidos como qualquer informação relacionada a uma pessoa física identificada ou identificável, como: nome, número de identificação, dados de localização, identificador online ou um ou mais fatores específicos à identidade física, fisiológica, genética, mental, econômica, cultural ou social dessa pessoa natural. Para evitar usos ilegais e não autorizados de dados pessoais, o GDPR exige esforços específicos das empresas e descreve os direitos dos consumidores sobre seus dados pessoais. O artigo publicado por Wieringa et al. (2019) traz, pois, as funções relacionadas à análise e privacidade de dados definidas pelo GDPR, tendo impacto nas práticas de LA.

Wachter (2018) inicia afirmando que o GDPR cria princípios governamentais de processamento de dados (GDPR, artigos 5 e 25) e que este estabelece novos padrões de proteção de dados relevantes para dispositivos de IOT. O autor também comenta que a literatura compara os antigos com os novos padrões do GDPR, para avaliar se a nova estrutura de proteção de dados é capaz de resolver a tensão inerente entre privacidade e identificabilidade na IOT. Esclarece, ainda, que o GDPR pode ajudar a aliviar alguns desses riscos, já que a proteção e a privacidade dos dados são as principais preocupações dos sistemas de IOT. Por fim, conclui que os controladores de dados devem tomar todas as medidas razoáveis para proteger a privacidade do usuário e, em particular, as exigidas pelo GDPR.

Tendo o último artigo selecionado junto à QP3, Yu e Zhao (2019) destacam que o objetivo do GDPR, que entrou em vigor em 25 de maio de 2018, é proteger a personalidade e a privacidade dos dados de um titular de dados. A proteção de dados envolve questões, como direitos civis, direitos de propriedade e responsabilidade contratual. A proteção de dados pessoais, para os autores, é o primeiro problema que precisa ser tratado no processo de utilização e transações de dados.

Da leitura dos artigos, é possível apresentar o conceito do GDPR, a legislação europeia que entrou em vigor em maio de 2018 com o objetivo de proteger a personalidade e a privacidade dos dados de um titular. Trata-se de um regulamento que fornece um conjunto de regras a serem seguidas por todos que utilizarem dados pessoais de europeus. Os artigos afirmam e esclarecem que a IA, os serviços e atividades realizadas em nuvem, as IOT, devem se encaixar na estrutura e seguir a legislação apresentada pelo GDPR, protegendo qualquer informação pessoal de europeus que possa ser usada para identificar direta ou indiretamente um indivíduo. É possível afirmar, ainda, que a própria estrutura do GDPR traz impacto nas práticas e no *design* de LA.

Por fim, apresenta-se a QP4, que buscava compreender como a aplicação do GDPR está contribuindo para a gestão universitária. Nessa questão, para sustentação da resposta, foram identificados (02) dois artigos, cuja análise discute-se a seguir.

Kitto e Knight (2019) apresentaram as Universidades The Open University e de Edimburgo desenvolvendo políticas de comunicação de coleta de dados pessoais com informações claras para os alunos e explicando quais os dados e os algoritmos que os analisam serão e/ou não usados. Essas políticas não só fornecem informações aos estudantes mas também orientação aos profissionais que constroem ferramentas de LA.

Hoel, Griffiths e Chen (2017) afirmam que se a coleta de dados, o armazenamento e o processamento de dados e a análise forem bem projetados, poderão contribuir para criar o perfil de aluno e entender melhor como os dados são usados em uma sociedade moderna.

Da leitura desses artigos, é possível analisar e compreender que o aluno é ponto central, sendo o GDPR a legislação que serve de orientação e controle para a utilização dos seus dados. Caso a gestão universitária deseje analisar os dados dos alunos, ela pode fazê-lo sem problemas. Contudo, para trabalhar com esses dados é preciso tratar os alunos com o devido respeito. Nesse sentido, as universidades The Open University e de Edimburgo explicam aos estudantes como os seus dados serão gerados e analisados, ou seja, a comunicação é clara, direta e aberta entre alunos e Universidade, de modo que não há dúvidas junto aos alunos sobre a coleta, a análise e como a divulgação dos dados irá ocorrer. Tudo isso corrobora a elaboração de LA, definindo onde e como serão realizadas as coletas e facilita os gestores em suas tomadas de ações.

4 CONSIDERAÇÕES FINAIS

Neste artigo, foram apresentados os resultados do mapeamento de uma RSL que permitiu identificar as pesquisas realizadas no período de 2015 até 2019 sobre a temática Learning Analytics e aplicação do General Data Protection Regulation (GDPR). A pesquisa ocorreu nas fontes de busca dos portais SCIENCE DIRECT, WEB OF SCIENCE e SCOPUS. O *string* de busca aplicado foi (Learning Analytics or Academic Analytics) e (General Data Protection Regulation or GDPR). Foram identificados (42) quarenta e dois artigos em um primeiro momento. Desse total, (10) dez textos foram selecionados e o restante, (32) trinta e dois, foi excluído. Sobre as (04) quatro QP que nortearam esta pesquisa, observou-se que todas foram atendidas, de modo que se alcança o objetivo inicialmente proposto.

A QP1, que trouxe os principais conceitos que estão sendo relacionados a Learning Analytics, identificou, inicialmente, que a ética vem influenciando os construtores de sistema de LA desde a concepção e que o Learning Analytics abre um cenário complexo de questões de privacidade e políticas, que, por sua vez, influenciam como os sistemas e práticas de análise de aprendizado são e serão projetados. As respostas obtidas para a QP2, questionando quais são as tecnologias de informação e comunicação (TICs) que estão sendo aplicadas à temática de Learning Analytics, foram da introdução das IA e da IOT como tecnologias relacionadas ao dia a dia do indivíduo. Também se fez menção à aplicação de sistemas em nuvem, como Amazon EC2 e Windows Azure, e a sistemas para gerenciamento e controle de acesso, como OpenID, PKI, SAML e Ucode. Quanto à QP3, que perguntou como o conceito de GDPR estava sendo aplicado nessas pesquisas, observou-se que os desenvolvedores de LA estão seguindo as orientações definidas pela legislação europeia. Além disso, evidenciou-se que o conceito permeia expressões-chave, como “proteção de dados pessoais”, “responsabilidade e transparência na utilização de dados pessoais”, “princípios governamentais de processamento de dados”, “padrões de proteção de dados”, “preservação da privacidade de forma justa e transparente”, “direito do indivíduo de saber como e o que de seus dados estão sendo utilizados”, bem como que o GDPR é qualquer informação que permita a identificação de um indivíduo. Quanto à QP4, a última e que trouxe como a aplicação da GDPR estava contribuindo para a gestão universitária, identificou-se a aplicação de políticas com informações claras por parte de universidades para os alunos sobre o tema de utilização de dados pessoais. Esse posicionamento colabora com a coleta e com o processamento dos dados, ajudando na formação do perfil do aluno e na gestão universitária de uma forma administrativa geral, além de ajudar aqueles que atuam na construção do LA.

Considerando os resultados colhidos, é importante destacar a contribuição teórica de dois artigos, em especial. O primeiro é o estudo de Kitto e Knight (2019), que respondeu às (04) quatro questões desta pesquisa. Posteriormente, o estudo de Hoel, Griffiths e Chen (2017), que deu sustentação para (03) três das questões propostas. Além disso, não se pode deixar de mencionar as próprias limitações desta pesquisa, clássicas em uma RSL. Uma situação a ser considerada é que estudos significativos para as áreas podem não ter sido integrados à seleção, assim como somente foram selecionados artigos “*open access*”.

Em síntese, é possível concluir com esta RSL que os estudos sobre o LA e o GDPR estão diretamente relacionados e que os temas merecem maior atenção por parte das empresas comerciais, das instituições governamentais, das universidades e dos pesquisadores. Existe um amplo espaço para investigações nessas áreas, como a realização de pesquisas acadêmicas e aplicação de projetos que possam ser direcionados para o *design* do LA tendo como orientação o GDPR, também para o controle de ações a serem tomadas para preservar os dados dos indivíduos na área comercial, assim como pontos a serem observados na construção do LA pelas universidades quanto à utilização dos dados de alunos.

REFERÊNCIAS

ALTMAN, I. **The environment and social behavior**: privacy, personal space, territory, and crowding. Monterey, California 93940: Brooks/Cole Publishing Company, 1975.

ANTUNES, N.; BALBY, L.; FIGUEIREDO, F.; LOURENCO, N.; MEIRA JR, W.; SANTOS, W. **Fairness and transparency of machine learning for trustworthy cloud services**. 2018, 48th Annual IEEE/IFIP INTERNATIONAL CONFERENCE ON DEPENDABLE SYSTEMS AND NETWORKS WORKSHOPS (DSN-W), 2018, p. 188-193, Doi: 10.1109/DSN-W.2018.00063.

BUTTERWORTH, M. The ICO and artificial intelligence: The role of fairness in the GDPR framework. **Computer Law & Security Review**, v.34, 2018, p.257-268.

CASTILHO, Carlos Albano Volkmer de. **O processo colaborativo na produção de informações**: gênese, sistemas e possíveis aplicações no jornalismo comunitário. 2009. Dissertação (Mestrado) - Universidade Federal de Santa Catarina. Florianópolis: UFSC, 2009.

CHATTI, M. A.; DYCKHOFF, A. L.; SCHROEDER, U.; THÜS, H. A. Reference model for learning analytics. **International Journal of Technology Enhanced Learning**, Geneva, Switzerland, v.4, n.5/6, p. 318-331, Jan. 2012.

CHI ZHOU, Amelie Chi et al. Privacy regulation aware process mapping in geo-distributed cloud data centers. **IEEE Transactions on Parallel and Distributed Systems**, v. 30, n. 8, p. 1872-1888, 2019.

DRESCH, A.; LACERDA, D. P.; ANTUNES JUNIOR, J. A. V. **Design Science Research**: método de pesquisa para avanço da ciência e tecnologia. Porto Alegre: Bookmann, 2015.

HACK, Josias Ricardo. **Introdução à educação a distância**. Florianópolis: LLV/CCE/UFSC, 2011.

HOEL, T.; CHEN, W. Data Sharing for Learning Analytics – designing conceptual artefacts and processes to foster interoperability. 24th INTERNATIONAL CONFERENCE ON COMPUTERS IN EDUCATION, 24th. **Proceedings...** India: Asia-Pacific Society for Computers in Education, 2016.

HOEL, Tore; GRIFFITHS, Dai; CHEN, Weiqin. The influence of data protection and privacy frameworks on the design of learning analytics systems. In: INTERNATIONAL LEARNING ANALYTICS & KNOWLEDGE CONFERENCE, 7., 2017. p. 243-252. **Proceedings of the...**New York, USA: ACM Press; 2017

JOHNSON, L.; SMITH, R.; WILLIS, H.; LEVINE, A.; HAYWOOD, K. **The NMC Horizon Report**: 2011 Higher Education Edition. New Media Consortium, 2011.

KITTO, K.; KNIGHT S. Practical ethics for building learning analytics. **British Journal of Educational Technology**, v.0, n.2019, p. 1-16. DOI: 10.1111/bjet.12868.

KITCHENHAM, B. A. **Guidelines for performing Systematic Literature Reviews in Software Engineering**. 2007.

KINDT, E. J. Why research may no longer be the same: about the territorial scope of the New Data Protection Regulation. **Computer Law & Security Review**, v.32, p.729-74, 2016.

KOEDINGER, Kenneth et al. An open repository and analysis tools for fine-grained, longitudinal learner data. In: EDUCATIONAL DATA MINING 2008: INTERNATIONAL CONFERENCE ON EDUCATIONAL DATA MINING, 1st., 2008. **Proceedings...**, 2008. p. 157-166.

MOISSA, B.; GASPARINI, I.; KEMCZINSKI, A. Educational data mining versus learning analytics: estamos reinventando a roda? Um mapeamento sistemático. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 26., 2015. **Anais...**, 2015. p.1167

SIEMENS, G.; BAKER, R. Learning analytics and educational data mining: towards communication and collaboration. In: Proceedings of the 2nd INTERNATIONAL CONFERENCE ON LEARNING ANALYTICS AND KNOWLEDGE, 2., 2012. **Anais...2012**. p. 252-254, 2012.

SIEMENS, G., LONG, P. Penetrating the fog: analytics in learning and education. **EDUCAUSE Review**, v.46, n.5, 2011.

SONI, G.; KODALI, R. A critical analysis of supply chain management content in empirical Research. **Business Process Management Journal**, v. 17, n. 2, p. 238-266. 2011.

THORPE, R. et al. Using knowledge within small and medium sized firms: a systematic review of the evidence. **International Journal of Management Reviews**, v. 7, n. 4, p. 257-281, 2005.

WACHTER, S. Normative challenges of identification in the Internet of Things: Privacy, profiling, discrimination, and the GDPR. **Computer Law & Security Review**, v.34, (2018), p.436-449.

WIERINGA, J.; MA, XIAO.; REUTTERER, T.; RISSELADA, H.; SKIERA, B. Data analytics in a privacy-concerned world. **Journal of Business Research**, <https://doi.org/10.1016/j.jbusres.2019.05.005> .

WILLS, D. B. The technology foresight activities of European Union data protection authorities. **Technological Forecasting & Social Change**, v.116 (2017), p.142-150.

YU, X.; ZHAO, Y. Dualism in data protection: balancing the right to personal data and the data property right. **Computer Law & Security Review**, v. 35, 105318, 2019

Revisão grammatical realizada por: Lovani Volmer

E-mail: lovaniv@feevale.br