

POLÍTICA EDUCACIONAL E CIÊNCIA DE DADOS NO BRASIL: UM CAMINHO POR CONSTRUIR

João Batista Carvalho Nunes
Universidade Estadual do Ceará
Brasil
joao.nunes@uece.br

Introdução

No Brasil, o uso de grandes bases de dados não é algo recente no campo da política educacional. Análises estatísticas têm sido realizadas sobre dados disponíveis em repositórios públicos ou produzidos por instituições de ensino e pesquisa ou organizações do terceiro setor.

Diante do crescimento exponencial na produção de dados digitais nas últimas décadas, seja por meio da ação humana, seja de forma automatizada, observou-se o surgimento e a rápida expansão da ciência de dados (*data science*) nas diversas áreas do conhecimento (PRIESTLEY; MCGRATH, 2019). Ela integra conhecimentos da estatística, computação e da área específica na qual está sendo aplicada (BLEI; SMYTH, 2017).

A ciência de dados tem se destacado por estar orientada à busca de soluções para problemas reais, assim como poder combinar dados digitais estruturados (armazenados em um banco de dados relacional) e não-estruturados (áudio, vídeo, postagens em redes sociais etc.). Sua crescente adoção na área de educação pode ser constatada, segundo McFarland *et al.* (2021), pela expressiva elevação de, no mínimo, cinco palavras-chave em publicações constantes na base ERIC de 2000 a 2020: *learning analytics* (análise da aprendizagem), *machine learning* (aprendizado de máquina), *artificial intelligence* (inteligência artificial), *data science* e *natural language processing* (processamento de linguagem natural). Inclusive, o destaque da análise da aprendizagem entre essas expressões pode ser observado no seu desenvolvimento na América Latina, incluindo o Brasil (CECHINEL *et al.*, 2020; NUNES, 2015). Em função de sua perspectiva abrangente, a ciência de dados pode ser considerada, ademais, um campo que inclui esses outros domínios.

Por outro lado, dada a amplitude dos objetos de pesquisa da política educacional, focados em “[...] ações ou omissões do Estado [...]” (SOARES, 2018, p. 11), a ciência de dados se mostra com maior poder de contribuição, demonstrando relevância na investigação de temas relacionados à política educacional (MCFARLAND *et al.*, 2021).

Este trabalho, por conseguinte, procura sintetizar os resultados de uma pesquisa de natureza exploratória, cujo objetivo é mapear as publicações que articulem política educacional e ciência de dados em âmbito nacional e internacional.

A metodologia foi assentada no paradigma pragmático e na abordagem quantitativa. Fez uso do método de mapeamento sistemático de publicações científicas (DERMEVAL; COELHO; BITTENCOURT, 2020), tomando, como fontes, as bases de resumo e citações: ERIC, principal base internacional na área de educação; e Scielo, principal base nacional de periódicos científicos (NUNES, 2021). A consulta nas bases compreende o período de 2000 até a data de sua realização (24/10/2022). Foram utilizadas na busca as seguintes expressões em inglês: “data science” AND “education policy”; “learning analytics” AND “education policy”; “machine learning” AND “education policy”; “artificial intelligence” AND “education policy”; e “natural language processing” AND “education policy”.

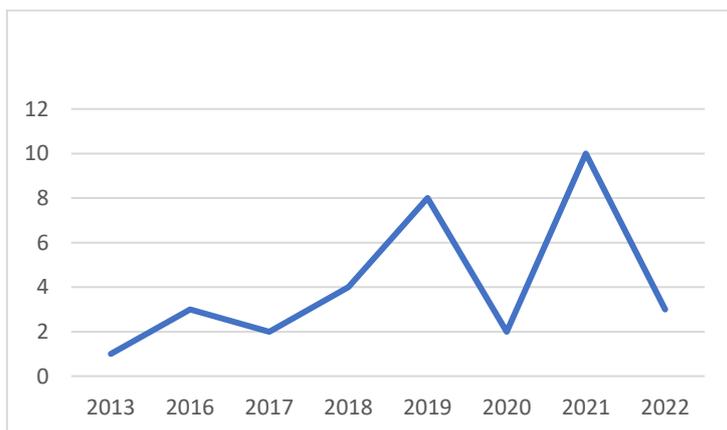
Desenvolvimento

Na base Scielo, ao se aplicar as expressões de busca explicitadas na seção anterior, nenhum registro de publicação foi obtido. É importante frisar que a base Scielo é multidisciplinar e inclui periódicos científicos de acesso aberto do Brasil, Argentina, Chile, Espanha, México, Portugal, África do Sul e de outros países da América Latina (NUNES, 2021).

Por outro lado, ao se realizar esse procedimento na base ERIC, obteve-se o total de 33 publicações, com início apenas em 2013, conforme Gráfico 1. Pode-se observar que as publicações mostram tendência de crescimento, com destaque para o ano de 2021, quando foram disponibilizadas 10 obras. Essas publicações estão distribuídas em 22 artigos, três livros (integrais ou coletâneas), um capítulo de livro, um livro de anais de evento e seis relatórios, publicados em países estrangeiros.

Embora tenham sido registradas apenas três publicações em 2022, a coleta desses dados foi realizada em 24 de outubro de 2022. Provavelmente outras publicações nesse ano ainda serão acrescentadas à base ERIC.

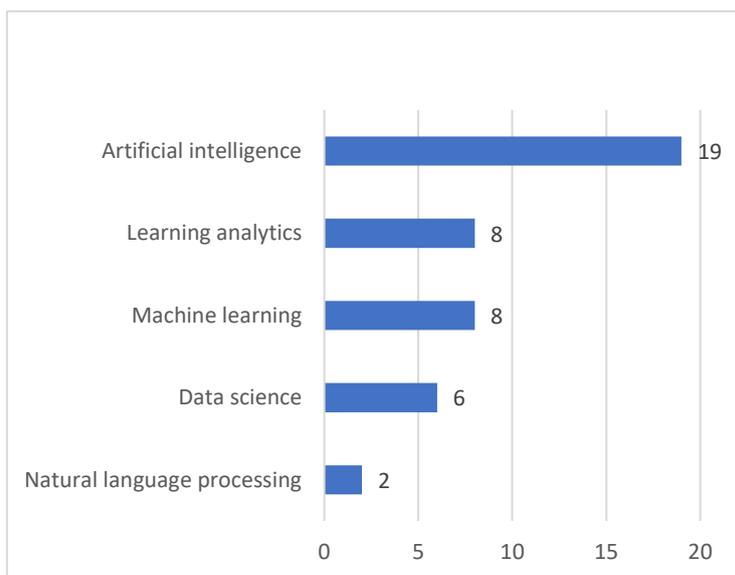
Gráfico 1 – Publicações na base ERIC relacionadas a “data science” e “education policy”, 2013-2022



Fonte: Elaboração própria.

A distribuição dessas 33 publicações pelos descritores “data science”, “learning analytics”, “machine learning”, “artificial intelligence” e “natural language processing” pode ser visualizada no Gráfico 2. Sobressaem-se as publicações cujo descritor é inteligência artificial, com 19 registros (57,6% do total). A expressão “data science” ainda é timidamente empregada, sendo responsável por somente 18,2% do total. O descritor “natural language processing” aparece com a menor frequência, representando apenas 6,1% do conjunto.

Gráfico 2 – Publicações na base ERIC, por descritores relacionados a “education policy”, 2013-2022



Fonte: Elaboração própria.

Nota: a soma das categorias não é 33, pois há publicações que aparecem em mais de um descritor.

O conjunto de 22 artigos obtidos está publicado em 16 periódicos internacionais, com destaque para o Journal of Education Policy (quatro artigos ou 18,2% desse total). Nos periódicos Education Policy Analysis, Critical Studies in Education e Assessment in Education: Principles, Policy & Practice, foram publicados dois (9,1%) artigos em cada. Nos demais periódicos, somente foi registrado um artigo (TABELA 1).

Tabela 1 – Artigos relacionados a “data science” e “education policy” na base ERIC, por periódicos, 2013-2022

Periódicos	Artigos
Advances in Learning and Behavioral Disabilities	1
Assessment in Education: Principles, Policy & Practice	2
Critical Studies in Education	2
Discourse: Studies in the Cultural Politics of Education	1
ECNU Review of Education	1
Education Policy Analysis	2
European Educational Research Journal	1
International Education Studies	1
International Journal of Artificial Intelligence in Education	1
Irish Educational Studies	1
Journal of Education Policy	4
Journal of Educational Methodology	1

Journal of Higher Education Policy and Management	1
Learning, Media and Technology	1
Research in Education	1
SoJo Journal: Educational Foundations and Social Justice Education	1
Total	22

Fonte: Elaboração própria.

Conclusões

Os resultados alcançados mediante a consulta às bases de resumos e citações ERIC e Scielo revelaram que ainda é incipiente o número de publicações científicas que articulam política educacional e ciência de dados. Observou-se a reduzida quantidade de obras que trazem a relação entre a expressão “education policy” e cada um dos seguintes descritores: “data science”, “learning analytics”, “machine learning”, “artificial intelligence” e “natural language processing”. No caso da Scielo, ainda não existe nenhum registro com alguma dessas combinações.

Considerando que os 16 periódicos, nos quais foram publicados os 22 artigos encontrados na base ERIC, são estrangeiros e que os demais 11 trabalhos também foram publicados fora do Brasil, pode-se afirmar que existe um promissor campo de pesquisa a ser construído no País, articulando a política educacional e a ciência de dados. Para tanto, propõe-se o investimento na formação em ciência de dados para os pesquisadores em educação, a fim de que, para além de parcerias que possam ser constituídas com profissionais das áreas de Ciências da Computação, Estatística ou afins, eles possam usar o conjunto de técnicas e ferramentas disponíveis nessa área para contribuir com a melhoria contínua do processo educacional.

Referências

- BLEI, David M.; SMYTH, Padhraic. Science and data science. **PNAS**, v. 114, n. 33, p. 8689-8692, 2017. DOI: 10.1073/pnas.1702076114.
- CECHINEL, Cristian; OCHOA, Xavier; SANTOS, Henrique Lemos dos; NUNES, João Batista Carvalho; RODÉS, Virginia; QUEIROGA, Emanuel Marques. Mapping learning analytics initiatives in Latin America. **British Journal of Educational Technology**, v. 51, n. 4, p. 892-914, 2020. DOI: 10.1111/bjet.12941.

DERMEVAL, Diego; COELHO, Jorge A. P. de M.; BITTENCOURT, Ig I. Mapeamento Sistemático e Revisão Sistemática da Literatura em Informática na Educação. In: JAQUES, Patrícia Augustin; SIQUEIRA; Sean; BITTENCOURT, Ig; PIMENTEL, Mariano (Org.). **Metodologia de Pesquisa Científica em Informática na Educação: abordagem quantitativa**. Porto Alegre: SBC, 2020. p. 1-26.

MCFARLAND, Daniel A.; KHANNA, Saurabh; DOMINGUE, Benjamin W.; PARDOS, Zachary A. Education data science: past, present, future. **AERA Open**, v. 7, n. 1, p. 1–12, 2021. DOI: 10.1177/2332858421105205.

NUNES, João Batista Carvalho. Estado da arte sobre analítica da aprendizagem na América Latina. In: LATIN AMERICAN WORKSHOP ON LEARNING ANALYTICS, 1, 2015, Maceió, 2015. **Anais...** Maceió: SBC, 2015. p. 1024-1033.

NUNES, João Batista Carvalho. Busca de literatura científica na área de Educação. In: NUNES, João Batista Carvalho; FARIAS, Isabel Maria Sabino; NÓBREGA-THERRIEN, Silvia Maria (Org.). **Pesquisa científica para iniciantes: retomando o debate**. Fortaleza: EdUECE, 2021. p. 11-23.

PRIESTLEY, Jennifer; MCGRATH, Robert. J. The evolution of data science: a new mode of knowledge production. **International Journal of Knowledge Management**, v. 15, n. 2, p. 1-13, 2019. DOI: 10.4018/IJKM.2019040106.

SOARES, Solange Toldo. A pesquisa em Política Educacional no Brasil: revisão de literatura. **Revista de Estudios Teóricos y Epistemológicos en Política Educativa**, v. 3, p. 1-17, 2018. DOI: 10.5212/retepe.v.3.013.