



ESTRATÉGIAS DE OBTENÇÃO E PROCESSAMENTO DE DADOS NA NOVA CIÊNCIA DAS CIDADES: SUBSÍDIOS PARA O MONITORAMENTO URBANO ATRAVÉS DE DADOS ABERTOS (SESSÃO TEMÁTICA)

Ana Lúcia Abrão Latrônico

Universidade Regional de Blumenau | analuciarq@gmail.com

Marcos Mattedi

Universidade Regional de Blumenau | mam@furb.br

Sessão Temática 01: Produção do Espaço Urbano e Regional

Resumo: O artigo examina o uso de dados abertos para análise, planejamento, gestão e desenho urbano. Para isso, utilizou uma metodologia mista de análise de dados, realizando análise bibliométrica de 115 artigos em nível macro e revisão integrativa de 25 artigos em nível micro. Os resultados indicam que o campo é multidisciplinar, multitemático e está em formação, com crescimento significativo nos últimos cinco anos. O OpenStreetMap destaca-se como principal base de dados abertos, sendo utilizado em 68% dos artigos analisados. As ferramentas mais empregadas para análise são GIS e linguagens de programação como Python e R, além de técnicas estatísticas e de machine learning. Os estudos concentram-se na escala da cidade e apresentam alta transferibilidade por utilizarem dados globalmente disponíveis. Conclui-se que dados abertos constituem uma alternativa viável para estudos urbanos, permitindo análises comparativas entre cidades e validação de métodos em diferentes contextos.

Palavras-chave: planejamento urbano; dados abertos; monitoramento urbano; nova ciência das cidades

STRATEGIES FOR OBTAINING AND PROCESSING DATA IN THE NEW SCIENCE OF CITIES: SUBSIDIES FOR URBAN MONITORING THROUGH OPEN DATA

Abstract: *This article examines the use of open data for urban analysis, planning, management, and design. To this end, it used a mixed methodology of data analysis, performing bibliometric analysis of 115 articles at the macro level and an integrative review of 25 articles at the micro level. The results indicate that the field is multidisciplinary, multithematic, and still in development, with significant growth in the last five years. OpenStreetMap stands out as the main open database, being used in 68% of the articles analyzed. The most commonly used tools for analysis are GIS and programming languages such as Python and R, in addition to statistical and machine learning techniques. The articles focus on the city scale and present high transferability because they use globally available data. The conclusion is that open data constitutes a viable alternative for urban studies, allowing comparative analyses between cities and validation of methods in different contexts.*

Keywords: *urban planning; open data; urban monitoring; new science of the cities*

ESTRATEGIAS PARA LA OBTENCIÓN Y PROCESAMIENTO DE DATOS EN LA NUEVA CIENCIA DE LAS CIUDADES: SUBVENCIONES PARA EL MONITOREO URBANO A TRAVÉS DE DATOS ABIERTOS

Resumen: *El artículo examina el uso de datos abiertos para el análisis, la planificación, la gestión y el diseño urbano. Para ello se utilizó una metodología de análisis de datos mixta, realizando un análisis bibliométrico de 115 artículos a nivel macro y una revisión integradora de 25 artículos a nivel micro. Los resultados indican que el campo es multidisciplinario, multitemático y en formación, con un crecimiento significativo en los últimos cinco años. OpenStreetMap destaca como la principal base de datos abierta, siendo utilizada en el 68% de los artículos analizados. Las herramientas más utilizadas para el análisis son los SIG y los lenguajes de programación como Python y R, así como técnicas estadísticas y de aprendizaje automático. Los artículos se centran en la escala de la ciudad y son altamente transferibles ya que utilizan datos disponibles a nivel mundial. La conclusión es que los datos abiertos constituyen una alternativa viable para los estudios urbanos, permitiendo análisis comparativos entre ciudades y validación de métodos en diferentes contextos.*

Palabras clave: *planificación urbana; datos abiertos; monitoreo urbano; nueva ciencia de las ciudades*

INTRODUÇÃO

As transformações na dinâmica urbana acontecem de maneira cada vez mais acelerada, representando um desafio para a gestão urbana. Neste sentido, a disseminação das tecnologias de informação e comunicação (TICs) representa tanto uma condicionante para essas mudanças quanto um artifício para lidar com elas. Isso acontece porque, mesmo que as transformações sejam cada vez mais rápidas, também nunca tivemos tantos dados para compreendê-las. Assim, diversos autores (Kandt; Batty, 2021; Kourtit; Nijkamp; Steenbruggen, 2017; Bibri, 2021) têm ressaltado a importância de um urbanismo orientado a dados, ressaltando sua importância também para alcançar aos Objetivos do Desenvolvimento Sustentável (ODS) (Jain; Espey, 2022). Neste contexto, Batty (2013) apresenta uma Nova Ciência das Cidades, onde propõe uma mudança de paradigma no seu estudo, mudando o foco para os fluxos (de pessoas, informações e recursos) e para as redes que conectam os elementos urbanos. Além disso, destaca o potencial do uso do *big data* e novas tecnologias para compreender as dinâmicas urbanas.

Complementando a ideia de Batty (2013), Ludwig (2022) sustenta que as transformações que ocorreram nas cidades no final do Século XX e início do Século XXI levaram a um Planejamento Urbano baseado no Paradigma da Informação. Isso significa que a cidade é considerada um sistema aberto, com uma diversidade de elementos que interagem de forma dinâmica e não linear. Por isso, o planejamento através do paradigma da informação aborda a cidade como um espaço que está sempre em transformação a partir de decisões de agentes urbanos, no qual os dados são essenciais para compreender a cidade (Ludwig, 2022). Assim, abordar a cidade através de dimensões temporais e espaciais torna-se essencial. Neste sentido, uma estratégia de monitoramento urbano que permita a compreensão e previsão destas dinâmicas urbanas é necessária, para que seja possível fazer sua gestão e assim minimizar possíveis problemas.

Um novo paradigma de planejamento urbano envolve também novas formas de obtenção, manipulação e compreensão dos dados. Estamos produzindo dados o tempo todo através de serviços de localização, câmeras, sensores, satélites, entre outras tecnologias. Estes dados possuem as dimensões temporais e espaciais necessárias para a compreensão das dinâmicas aceleradas que as cidades experimentam, portanto, devem ser utilizados para estudar as cidades. Além disso, muitos destes dados estão disponíveis de forma aberta na internet, ampliando seu acesso. Estes dados abertos incluem desde dados censitários, produzidos pelo poder governamental; até dados extraídos de imagens de satélite e plataformas alimentadas de forma colaborativa. Por isso, é necessário compreender as possibilidades destas bases de dados em um contexto de monitoramento. Assim, a pergunta de pesquisa deste artigo é *como dados abertos podem ser utilizados para análise, planejamento, gestão e desenho urbano?*

Os dados abertos podem ser uma fonte alternativa de baixo custo para que governos compreendam as dinâmicas de seus territórios. No entanto, segundo Jain e Espey (2022), mesmo que os dados abertos possibilitem que o poder governamental os utilize para

monitorar e influenciar a tomada de decisões, ainda existem muitas barreiras para sua adoção. Isso pode acontecer pela própria falta de conhecimento da multiplicidade de bases de dados abertas. Como Chakraborty et al. (2015) destacam, gerentes e planejadores urbanos precisam de orientações mais detalhadas sobre como navegar em um campo extenso de conjuntos de dados, além de ferramentas e técnicas variadas. Por isso, o entendimento de como estes dados podem ser utilizados é de fundamental importância, tanto para a comunidade acadêmica quanto para gestores urbanos. Assim, este artigo pode dar subsídios para o uso de dados abertos para o estudo das cidades.

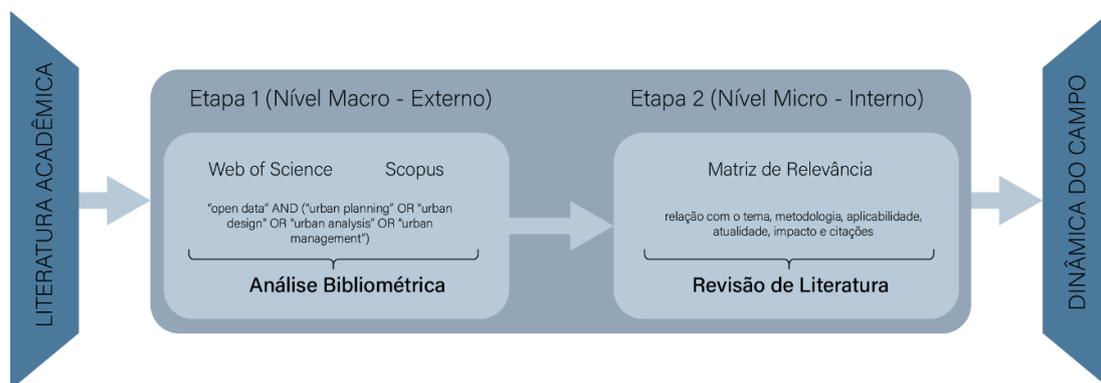
Assim, este artigo possui um objetivo geral e quatro objetivos específicos, sendo o Objetivo Geral *analisar como dados abertos estão sendo utilizados para análise, planejamento, gestão e desenho urbano*. Os objetivos específicos são: a) Compreender o campo de pesquisa que envolve o uso de dados abertos para análise, planejamento, gestão e desenho urbano; b) Conceber quais as características internas dos trabalhos que utilizam dados abertos para o estudo das cidades; c) Identificar as bases de dados abertos utilizadas para análise, planejamento, gestão e desenho urbano; e d) Apontar quais as ferramentas e técnicas utilizadas para análise urbana utilizando dados abertos.

Para atingir os objetivos propostos, uma estratégia mista de análise de dados foi elaborada. Primeiramente, a seleção dos dados foi realizada em bases de dados de literatura acadêmica em dois níveis: 1) Nível Macro e 2) Nível Micro. Em seguida, para o Nível Macro, efetuamos a análise bibliométrica para compreensão do campo de pesquisa que envolve o uso de dados abertos para o estudo das cidades. Para o Nível Micro, analisamos os artigos internamente, para compreender as características dos trabalhos publicados, identificar as bases de dados abertas utilizadas, bem como as ferramentas utilizadas para a manipulação. Assim, o artigo está estruturado da seguinte forma: a Seção 2 detalha a seleção dos dados, a Seção 2.1 trata da análise no Nível Macro, a Seção 2.2 aborda a análise no Nível Micro, e a Seção 3 conclui o artigo.

SELEÇÃO DOS DADOS

A reconstrução do padrão de utilização de dados abertos na análise, planejamento e gestão urbana envolve, primeiramente, a compreensão da dinâmica do campo de pesquisa. A descrição da dinâmica do campo pode ser efetuada através da revisão de literatura acadêmica. A revisão de literatura acadêmica possibilita a síntese dos resultados de pesquisas anteriores, o que permite compreender a dinâmica de um campo de pesquisa (Zupic; Carter, 2015). Neste sentido, para Callon, Law e Rip (1986), o problema de conceber a dinâmica da ciência e da tecnologia exige o uso de métodos quantitativos, mas estes são melhor vistos como um meio para buscar o qualitativo. Assim, elaboramos uma estratégia mista para revisão do campo, baseada na utilização de técnicas quantitativas e qualitativas. Esta estratégia consiste em, primeiramente, selecionar os artigos para a criação da base de dados. Este processo é dividido em dois níveis: 1) Nível Macro e 2) Nível Micro (Figura 1).

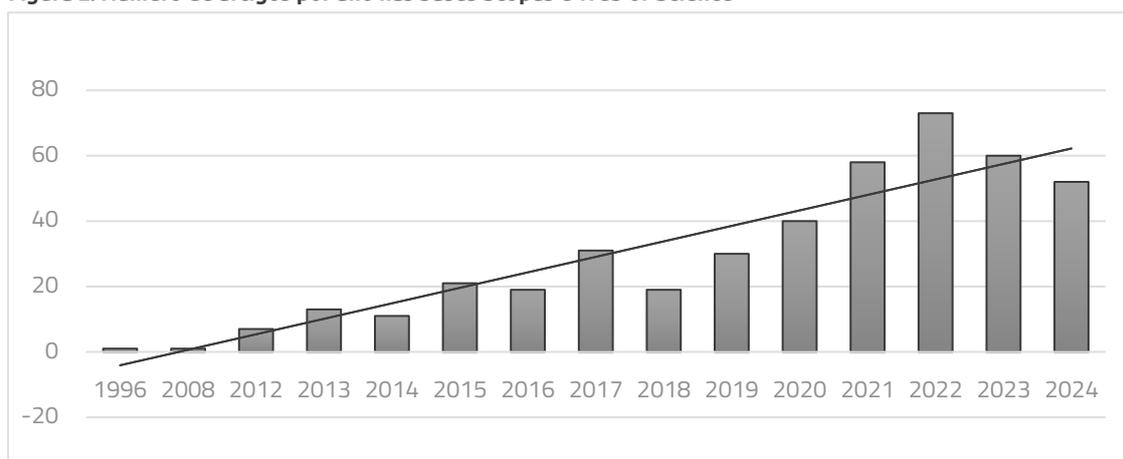
Figura 1: diagrama de seleção dos dados.



Fonte: os autores.

A seleção no Nível Macro baseou-se em critérios gerais de seleção dos artigos. Esta etapa começou pela escolha da fórmula de pesquisa e das bases de dados a serem utilizadas. As bases utilizadas foram a Scopus e a Web of Science, por serem as plataformas mais conhecidas e com maior abrangência (Roemer; Borchardt, 2015). Após a busca pela fórmula de pesquisa, os critérios de seleção foram: 1) Idioma: foram selecionadas apenas publicações em inglês, por ser a linguagem com o maior número de artigos; 2) Ano de publicação: publicações feitas nos últimos cinco anos (2020-2024). Este critério de seleção foi aplicado porque foi a partir de 2020 que o número de publicações começou a subir (Figura 2). Além disso, presume-se que as inovações exploradas anteriormente estão incorporadas nas publicações mais recentes; 3) Tipo de publicação: apenas artigos revisados por pares. Após realizadas as buscas nas duas bases de dados, os artigos repetidos foram removidos (Quadro 1).

Figura 2: Número de artigos por ano nas bases Scopus e Web of Science



Nota: artigos publicados até Agosto de 2024.

Fonte: os autores - com base nos dados da Scopus e Web of Science.

Quadro 1: Fórmula de busca e critérios de seleção no Nível Macro

Fórmula de busca					
Campos	String de busca				
Title, Abstract, Keywords	"open data" AND ("urban planning" OR "urban design" OR "urban analysis" OR "urban management")				
Filtros					
Base de dados	Busca geral	Filtro 1: idioma – inglês	Filtro 2: ano 2020 a 2024	Filtro 3: tipo – artigos	Remoção repetidos
Scopus	301	285	184	108	
Web of Science	135	130	85	67	115
Total	436	415	269	175	

Fonte: os autores.

A seleção em Nível Micro buscou determinar a relevância dos artigos para construção da amostra que será analisada através de revisão de literatura. A relevância dos artigos foi definida, primeiramente, através do número de citações de cada publicação. Além do número de citações, para evitar o viés de anterioridade, considerou-se também o fator de impacto das revistas. Isso foi necessário pois como a amostra considera artigos publicados nos últimos cinco anos, publicações recentes podem ser altamente relevantes, mas não apresentarem citações ainda. Após esta etapa, a amostra apresentava 54 artigos (Quadro 2). Para refinar a seleção, foi elaborada uma Matriz de Relevância que determinou critérios para pontuação dos artigos com base na leitura do seu resumo. Os critérios foram: relação com o tema, metodologia, aplicabilidade, atualidade, impacto e citações. A amostra final para o Nível Micro conta com 25 artigos que serão analisados internamente.

Quadro 2: Critérios de seleção inicial em Nível Micro.

Amostra inicial – Nível Macro	Critério 1: citações – mínimo 20	Critério 2: fator de impacto – mínimo 6	Remoção repetidos
115 artigos	26 artigos	36 artigos	54 artigos
	62 artigos		

Fonte: os autores.

Feita a seleção, foi realizada a análise dos dados em Nível Macro e Nível Micro. Para o Nível Macro, utilizamos a Análise Bibliométrica para examinar os 115 artigos. Este método teve como objetivo a compreensão do campo de pesquisa que envolve o uso de dados abertos para o estudo das cidades. Assim, foram analisados aspectos relativos às revistas acadêmicas, autores, países, palavras-chave e a relação entre eles. Para o Nível Micro, o método utilizado foi o de Revisão Integrativa de Literatura. Neste caso o objetivo foi a análise interna dos artigos selecionados, através de sua leitura completa. Dessa forma, foram consideradas na investigação as características dos artigos, incluindo escala de análise, abordagem do estudo, conceitos relacionados a Planejamento Urbano utilizados, as bases de dados abertas utilizadas, as ferramentas de análise e o nível de transferibilidade do estudo. Estas análises permitiram compreender tanto o campo de pesquisa quanto as características individuais dos estudos.

NÍVEL MACRO – EXTERNO

A bibliometria é frequentemente compreendida como um conjunto de métodos quantitativos usados para medir, rastrear e analisar a literatura acadêmica (Roemer; Borchardt, 2015). Este conjunto de métodos busca avaliar a performance das revistas, autores, instituições, publicações, entre outros. No entanto, é possível introduzir análises qualitativas que permitem mapear a dinâmica de um campo de pesquisa (Noyons; Moed; Luwel, 1999). Neste sentido, o mapeamento da dinâmica da ciência permite descrever a estrutura intelectual de um campo, ajuda na busca de informações relevantes e fornece uma base empírica para testar conceitos abstratos (De Bellis, 2009). Por isto, nesta análise bibliométrica combinamos métodos quantitativos e qualitativos para compreender o desenvolvimento do problema proposto. Esta análise busca compreender aspectos relativos às revistas, artigos, autores e aos temas abordados pelas publicações selecionadas.

Um elemento relevante para a compreensão da estrutura intelectual de um campo é a revista científica. A análise de periódicos materializa a atividade científica tanto nas suas dimensões institucionais quanto cognitivas. Neste sentido, permite compreender as áreas do conhecimento em que se estabelece o tema de pesquisa. Por isso, sua análise é importante para a compreensão do desenvolvimento da dinâmica do campo (Mattedi; Spiess, 2023). Desta forma, os 115 artigos da base de dados encontram-se publicados em 74 periódicos. Sobre estes periódicos, considerou-se os atributos: *fator de impacto, com base no Journal Citation Reports – JCR (FI)*; *áreas do conhecimento, também com base nas categorias do JCR (AC)*; *o número de artigos da base de dados publicados em cada revista (NA)*; e *o total de citações dos artigos da base de dados em cada uma das revistas (TC)* (Quadro 3).

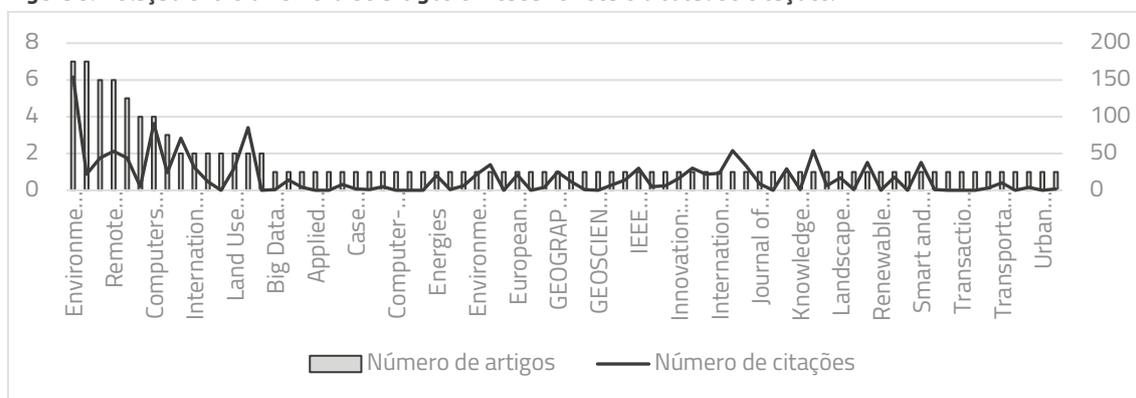
Quadro 3: Relação de periódicos com 2 ou mais artigos publicados na base de dados.

Revista	FI	AC	NA	TC
Sustainability (Switzerland)	3,3	environmental sciences environmental studies green & sustainable science & technology	7	22
Environment and Planning B: Urban Analytics and City Science	2,6	regional & urban planning environmental studies urban studies geography	7	154
Remote Sensing	4,2	imaging science & photographic technology environmental sciences geosciences, multidisciplinary remote sensing	6	53
ISPRS International Journal of Geo-Information	2,8	geography, physical computer science, information systems remote sensing	6	44
Sustainable Cities and Society	10,5	energy & fuels construction & building technology green & sustainable science & technology	5	44
Computers, Environment and Urban Systems	7,1	regional & urban planning environmental studies geography	4	91
Cities	6	urban studies	4	5
Frontiers in Public Health	3	public, environmental & occupational health	3	24
International Journal of Digital Earth	3,7	geography, physical remote sensing	2	31
Applied Geography	4	geography	2	71
Journal of Urban Planning and Development	1,7	regional & urban planning engineering, civil urban studies	2	0
Land Use Policy	6	environmental studies	2	31
International Journal of Environmental Research and Public Health	4,614	public, environmental & occupational health environmental sciences	2	12
Transportation Research Record	1,6	engineering, civil transportation science & technology	2	0
Remote Sensing of Environment	11,1	imaging science & photographic technology environmental sciences remote sensing water resources	2	85

Fonte: os autores.

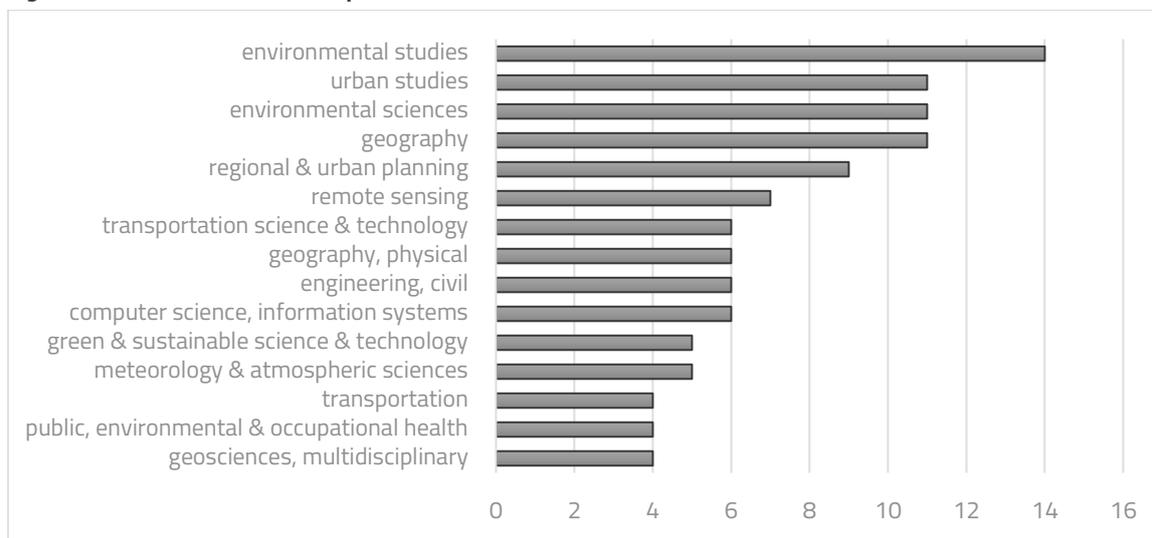
Observa-se que há uma concentração de quase 50% dos artigos da base de dados em apenas 15 revistas, enquanto as 59 revistas restantes apresentam apenas um artigo. Porém, não se constata, necessariamente, uma relação direta entre o número de artigos publicados (NA) em cada revista e o total de citações (TC) (Figura 3). Em relação às áreas do conhecimento (AC), as revistas encontram-se em um conjunto de 44 categorias diferentes. As categorias com maior predominância são as que envolvem estudos ambientais, geografia e estudos urbanos (Figura 4). Destacam-se também categorias que envolvem tecnologias, como sensoriamento remoto e ciências da computação. Além disso, categorias com temas mais específicos aparecem, como transportes, engenharia civil, sustentabilidade, meteorologia e saúde. Isso mostra que a dinâmica do campo de pesquisa envolvendo dados abertos e planejamento urbano é multidisciplinar, podendo ser abordado por diversos pontos de vista.

Figura 3: Relação entre o número de artigos em cada revista e o total de citações.



Fonte: os autores.

Figura 4: Áreas do conhecimento predominantes.

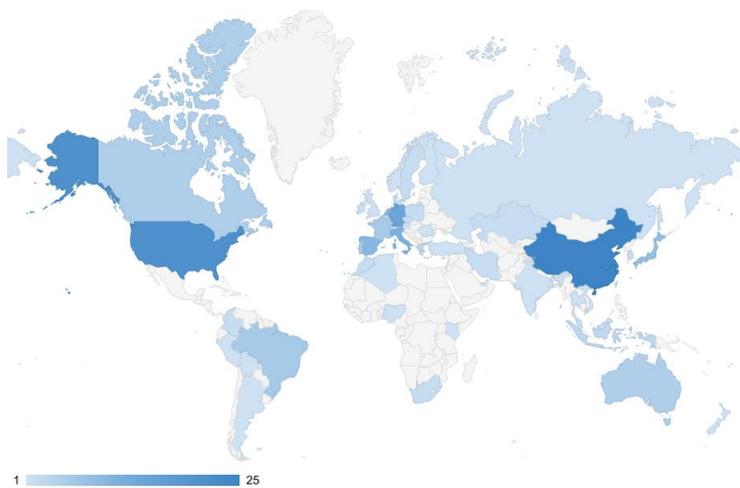


Fonte: os autores.

A atividade científica não opera desvinculada do meio em que o pesquisador está inserido. Por isso, pode-se dizer que a dinâmica de um campo de pesquisa também possui uma dimensão territorial. Esta dimensão está relacionada com os países e instituições mais produtivas em termos de pesquisa. Assim, os artigos da base de dados são procedentes de

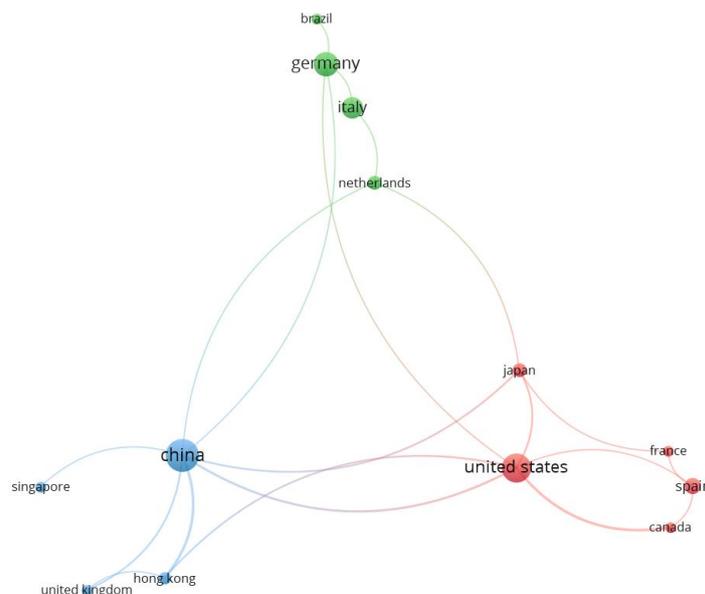
51 países, sendo os mais produtivos: China, Estados Unidos, Alemanha, Itália e Espanha (Figura 5). Com base na análise de co-autoria de artigos utilizando como unidade os países, elaborada com o auxílio do software *VOSviewer*, também identificamos as redes de colaboração em que estes países se inserem. Conforme a Figura 6, foram identificados três *clusters* de colaboração: 1) Países asiáticos e Reino Unido, sendo a China a principal representante; 2) Estados Unidos, Canadá, Japão e Europa, sendo os EUA o principal representante; 3) Países Europeus e Brasil, sendo a Alemanha o principal representante. Não se utilizou como unidade de análise as instituições porque identificou-se que havia poucas instituições em comum.

Figura 5: Países mais produtivos



Fonte: os autores.

Figura 6: Redes de colaboração entre os países

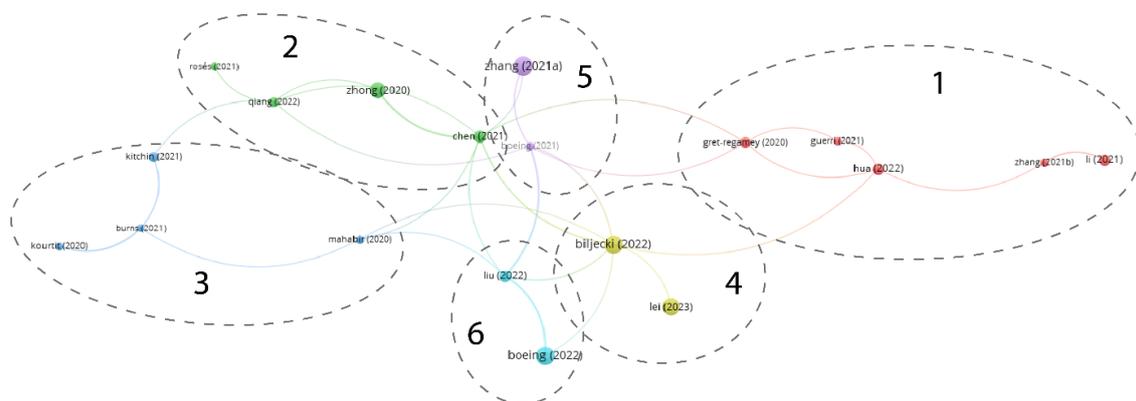


Fonte: os autores.

Para mapear um campo de pesquisa através da similaridade entre artigos, vários métodos podem ser utilizados. Entre eles, destacam-se: a) *Análise de Co-citações* (Small, 1973); b) *Acoplamento Bibliográfico* (Kessler, 1963). Enquanto a análise de Co-citações conecta artigos com base na ocorrência conjunta em listas de referências, o Acoplamento Bibliográfico encontra a similaridade entre publicações baseado no número de referências que elas compartilham. Neste sentido, a análise de Co-citações carrega mais informações para documentos altamente citados, já que a conexão entre artigos é estabelecida através dos autores que estão citando os trabalhos examinados (Zupic; Carter, 2015). Por isso, o Acoplamento Bibliográfico é mais indicado para o mapeamento de artigos atuais (Small, 1999), uma vez que pode contornar o viés de anterioridade (artigos antigos contém mais citações). Assim, o Acoplamento Bibliográfico permite encontrar referências chave para o campo de pesquisa, bem como identificar nichos de pesquisa dentro do campo.

Portanto, o Acoplamento Bibliográfico foi utilizado para encontrar a similaridade entre os 26 artigos mais citados (artigos com no mínimo 20 citações). Entre estes artigos, encontrou-se referências compartilhadas entre 19 deles, possibilitando a formação de 6 clusters (Figura 7). É possível identificar assuntos em comum para cada cluster, proporcionando o encontro de nichos específicos dentro do campo de pesquisa. A maioria das referências compartilhadas em cada um dos clusters apresentam artigos de periódicos similares aos estudos realizados pelos trabalhos analisados. Não foi possível identificar uma publicação principal que apresente uma “base teórica” entre todos os artigos. No entanto, publicações clássicas da área de Planejamento Urbano são encontradas, como *The Image of de City* de Kevin Lynch e *The City in History* de Lewis Mumford. Isso pode indicar que o campo de pesquisa possui uma característica majoritariamente técnica, que utiliza estudos experimentais para testar argumentos pré-estabelecidos. As características de cada cluster encontram-se no Quadro 4.

Figura 7: grafo de acoplamento bibliográfico.



Fonte: os autores.

Quadro 4: Descrição dos clusters de acoplamento bibliográfico.

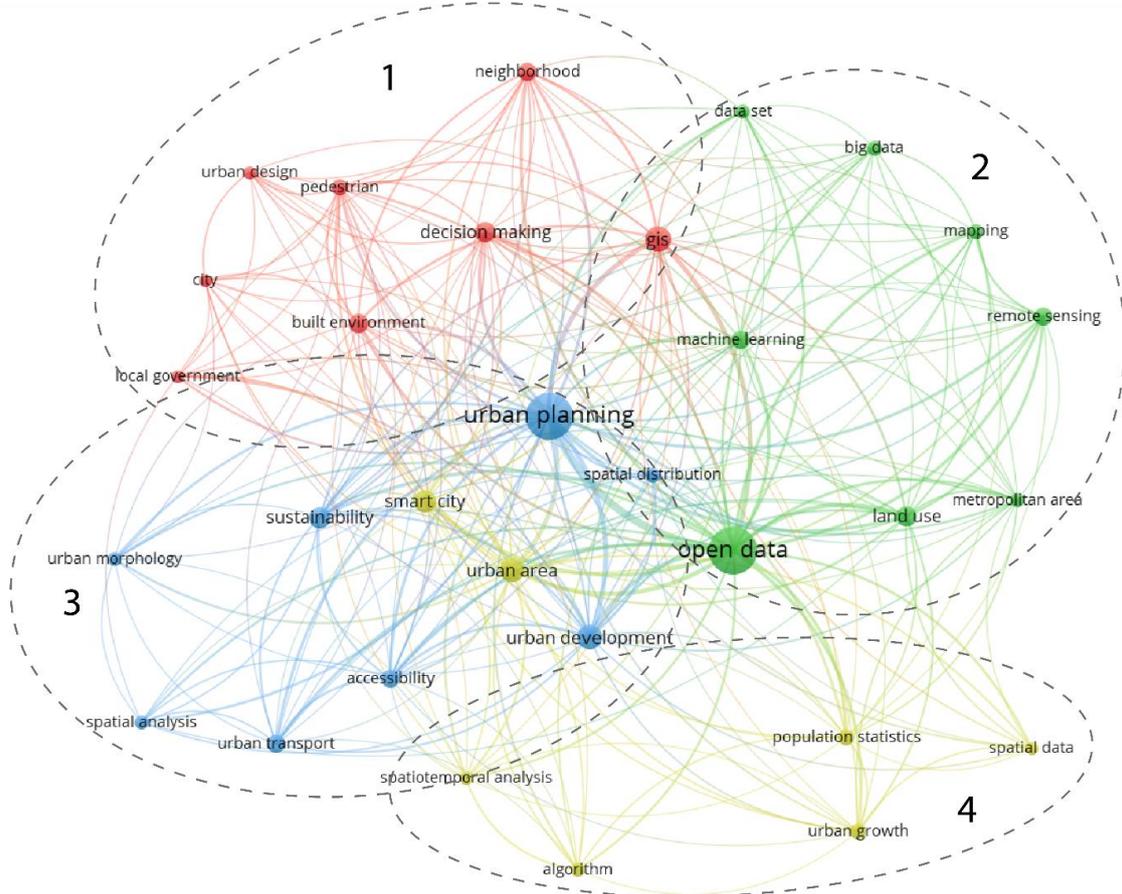
Cluster	Artigos	Características do Cluster
1	Gret-Regamey, A. et al. (2020) Guerrini, G. et al. (2021) Hua, J. et al. (2022) Li, H. et al. (2021) Zhang, P. et al. (2021)	Este cluster apresenta dois assuntos principais: 1) A relação entre o urbano e fatores ambientais; e 2) A relação entre preço e uso do solo. Alguns dos temas-chave incluem: o impacto da densidade urbana nos serviços ecossistêmicos; análise de ilhas de calor urbanas; vegetação em espaços urbanos; preços de habitação e sua relação com o uso do solo do entorno.
2	Chen, B. et al. (2021) Qiang, D. et al. (2022) Rosés, R. et al. (2021) Zhong, Y. et al. (2020)	O enfoque deste cluster encontra-se na análise do uso do solo. Alguns dos temas abordados incluem: o desenvolvimento de métodos para a classificação de uso do solo; a relação entre modos de transporte, uso do solo e desenvolvimento urbano; a busca de padrões de crime no ambiente urbano.
3	Burns, R. et al. (2021) Kitchin, R. et al. (2021) Kourtit, K. et al. (2020) Mahabir, R. et al. (2020)	Os artigos deste cluster apresentam uma perspectiva crítica sobre Smart Cities, governança baseada em dados, suas implicações sociais e desafios. Entre os temas tratados, encontram-se: análise crítica de projetos de Smart Cities; governança no contexto de Smart Cities; considerações sobre sustentabilidade no contexto de Smart Cities; desigualdade social e uso de tecnologias de mapeamento.
4	Biljecki, F. et al. (2022) Lei, B. et al. (2023)	Este cluster aborda a morfologia das edificações das cidades. Alguns dos temas-chave apresentados são: desenvolvimento de indicadores para a análise da forma urbana; modelos tridimensionais de cidades e avaliação comparativa de cidades.
5	Boeing, G. (2021) Zhang, A. (2021)	O tema principal deste cluster é morfologia urbana. Dentro deste tema, os assuntos abordados são: análise da morfologia de redes viárias; padrões de desenho de ruas vistos de uma perspectiva histórica; vitalidade urbana na escala de quarteirão e análise da relação entre morfologia urbana e dependência de veículos automotores.
6	Boeing, G. et al. (2022) Liu, S. et al. (2022)	Os artigos deste cluster abordam a criação e o uso de softwares open source para a medição de indicadores urbanos. Os temas explorados incluem: acessibilidade urbana para pedestres; cidades saudáveis e sustentáveis; criação de ferramentas que possibilitam estudos replicáveis globalmente e a aplicação delas.

Fonte: os autores.

Para aprimorar a compreensão dos principais temas abordados pelos artigos utilizou-se a técnica de co-ocorrência de palavras-chave (Callon et al., 1983). Esta técnica permite o estabelecimento de um mapa conceitual do campo, identificando as relações entre as palavras-chave utilizadas nos artigos selecionados. O conceito por trás desta técnica é que, quando palavras aparecem frequentemente nos documentos, significa que os conceitos por trás destas palavras estão relacionados. A análise resulta em uma rede de temas e suas relações, representando assim o campo de pesquisa (Zupic; Carter, 2015). Assim, o grafo da Figura 8 demonstra a rede de palavras-chave com maior ocorrência nos artigos selecionados.

Para a criação do grafo, foram consideradas as palavras-chave do autor e as indexadas pelas bases de dados. O grafo está dividido em 4 clusters diferentes, que agrupam as palavras-chave relacionadas entre si, mostrando diferentes enfoques no campo de pesquisa. As características de cada Cluster encontram-se no Quadro 5.

Figura 8: grafo de co-ocorrência de palavras-chave.



Fonte: os autores.

Quadro 5: Descrição dos clusters da análise de co-ocorrência de palavras-chave.

Cluster	Keywords	Características do Cluster
1 (vermelho)	Built environment City Decision Making GIS Local Government Neighborhood Pedestrian Urban Design	Este cluster apresenta uma concentração de palavras que focam em aspectos sobre o ambiente físico das cidades (Built environment, City, urban design) e à gestão urbana (Decision making, local government). O cluster enfatiza o estudo das cidades na escala de bairro (neighborhood, pedestrian). Além disso, destaca a tecnologia GIS (Geographic Information System) como ferramenta utilizada para os estudos realizados sobre os temas deste cluster. Desta forma, a partir da análise do grafo deste cluster, pode-se identificar uma concentração de estudos sobre a cidade numa escala de bairro, utilizando GIS como ferramenta de trabalho.
2 (verde)	Big data Data set Land Use Machine Learning Mapping Metropolitan area Open data Remote Sensing	Tem como tema principal dados e tecnologias utilizadas para o planejamento urbano (big data, data set, machine learning, mapping, open data, remote sensing). Também sugere a escala (metropolitan área) e o tipo de análise que pode ser feita (land use). Apesar de não fazer parte do cluster, a palavra-chave "GIS" também apresenta proximidade, indicando uma relação com o grupo. Assim, a análise deste cluster indica os tipos de tecnologia que têm sido utilizados para coleta, processamento e visualização de dados espaciais e seu uso para questões urbanas.
3 (azul)	Accessibility Spatial analysis Spatial distribution Sustainability Urban development Urban morphology Urban planning Urban transport	Este cluster tem como tema central planejamento urbano (urban development, urban planning), mostrando um enfoque no aspecto físico das cidades (spatial analysis, spatial distribution, urban morphology). Além disso, inclui temas específicos abordados neste contexto (accessibility, sustainability, urban transport). A análise deste cluster sugere aspectos mais "tradicionais" no contexto do planejamento urbano, relacionando temas há muito tempo trabalhados dentro do campo de pesquisa.
4 (amarelo)	Algorithm Population statistics Smart city Spatial data Spatiotemporal analysis Urban area Urban Growth	Este cluster apresenta uma intersecção entre crescimento urbano (population statistics, urban growth), análise de dados espaciais (spatial data, spatiotemporal analysis), escala urbana (urban area, smart city) e tecnologia (algorithm). O termo Smart City evidencia uma conexão entre o crescimento urbano, o estudo das cidades e as novas tecnologias, funcionando como um elo para o cluster. No grafo gerado, os termos "smart city" e "urban area" também apresentam proximidade ao cluster 3, indicando uma forte relação delas com os temas abordados pelo agrupamento.

Fonte: os autores.

Com base nas técnicas utilizadas para o reconhecimento do campo de pesquisa (análise das revistas, análise territorial, acoplamento bibliográfico e co-ocorrência de palavras-chave), pode-se afirmar que: *a) É um campo multidisciplinar*: as diversas áreas do conhecimento das publicações e os temas identificados indicam que o campo envolve muitas disciplinas diferentes; *b) É um campo multitemático*: através do Acoplamento Bibliográfico e da Co-ocorrência de palavras-chave foi possível identificar diversos temas trabalhados dentro do campo. Também foi possível verificar que a maioria dos temas tratados já estão estabelecidos

no campo de Planejamento Urbano – porém, são introduzidas novas abordagens que utilizam tecnologias recentes como ferramenta para explorar estes temas; e *c) É um campo em formação*: o número de publicações no período analisado, a quantidade de citações e a desconcentração de publicações entre os pesquisadores indicam que o campo ainda está se formando.

NÍVEL MICRO – ANÁLISE INTERNA

No Nível Micro, o objetivo foi analisar internamente os artigos selecionados através da matriz de relevância. Para isso, o método escolhido foi o de Revisão Integrativa de Literatura. Revisões de literatura são úteis quando o objetivo é fornecer uma visão geral de um determinado de uma questão ou problema de pesquisa, ou mesmo analisar criticamente o tópico abordado (Snyder, 2019). Mais especificamente, uma Revisão Integrativa de Literatura procura integrar diferentes perspectivas sobre um tema, e por isso é uma boa opção para campos multidisciplinares (Steil, 2021). Além disso, Revisões Integrativas de Literatura são indicadas para tópicos emergentes, pois levam a uma conceituação inicial do tema (Torraco, 2005). Dessa forma, os artigos selecionados foram analisados procurando responder as questões relacionadas abaixo:

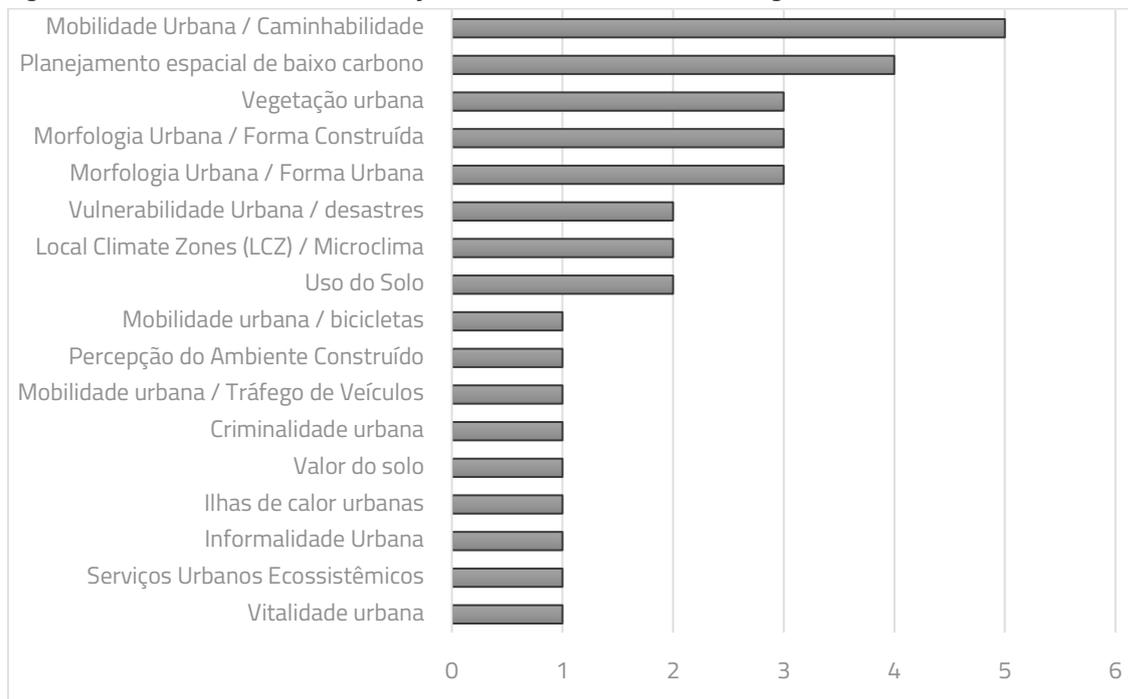
- a) Qual a escala de análise do artigo?
- b) Qual a abordagem do estudo, empírico ou teórico?
- c) Qual(is) o(s) principal(is) conceito(s) de Planejamento Urbano abordado(s) pelo artigo?
- d) Quais bases de dados abertas foram utilizadas?
- e) Quais foram as ferramentas e/ou técnicas utilizadas para a análise?
- f) Qual o nível de transferibilidade do método utilizado?

Os artigos explorados apresentam análises em diversos níveis, desde a escala de bairro até análises mais abrangentes a nível de país. No entanto, a maioria dos trabalhos expressa estudos na escala de cidade: dos 25 artigos selecionados, 21 deles analisam cidades. Destes 21 artigos, 12 deles investigam apenas uma cidade, e os outros 9 incluem duas ou mais cidades na pesquisa. Os artigos com duas ou mais cidades tratam a pesquisa de duas formas: 1) comparação de atributos entre as cidades; ou 2) validação do método proposto em contextos diferentes. Neste sentido, a avaliação comparativa entre cidades é um dos potenciais do uso de dados abertos para a análise urbana, considerando que os mesmos dados estão disponíveis para diversos locais diferentes. Da mesma forma, a proposta de um método de análise baseado em dados abertos favorece a aplicação do método por outros pesquisadores.

Todos os artigos selecionados apresentam uma abordagem empírica e/ou técnica. Isso se explica pelo próprio enfoque da seleção dos artigos, que coloca os dados como elemento

central da pesquisa. Ainda assim, os trabalhos utilizam conceitos e teorias já estabelecidos do campo de Planejamento Urbano como base para definição dos parâmetros de análise dos dados, ou mesmo para testar as hipóteses da teoria através dos dados. Dessa forma, o conceito mais utilizado pelos artigos selecionados foi o de Caminhabilidade (frequentemente associado com conceito de Cidade de 15 Minutos), sendo o enfoque de 5 dos artigos analisados. Outro conceito de destaque é o Planejamento Espacial de Baixo Carbono, com 4 estudos. Os tópicos Vegetação Urbana e Morfologia Urbana (relativa tanto à forma urbana quanto à forma construída) também foram muito empregados, sendo destaque em 3 artigos cada. Os principais temas dos artigos selecionados encontram-se na Figura 9.

Figura 9: Conceitos relacionados a Planejamento Urbano utilizados nos artigos selecionados.



Fonte: os autores.

Em vários dos artigos selecionados, os conceitos relacionados a Planejamento Urbano foram utilizados para a criação de indicadores. Os estudos de Boeing et al. (2022), Biljecki e Chow (2022), Hua et al. (2022), Zhang, A. et al. (2021), Grêt-Regamey et al. (2020), Mahabir et al. (2020), Liu et al. (2021), Wu et al. (2023) e Shartova, Mironova e Grischenko (2023) usam dados abertos para criar indicadores relativos aos temas abordados pelos trabalhos, de forma a quantificar determinados atributos das cidades analisadas. O uso de indicadores permite a avaliação comparativa entre cidades diferentes, além da comparação da mudança de características através do tempo em uma mesma cidade. Por isso, a criação de indicadores que podem ser interpretados através de dados abertos é uma iniciativa valiosa para o planejamento e gestão urbana. Da mesma forma, o uso de dados abertos permite que diferentes atores da sociedade façam as análises, além de possibilitar a replicabilidade do indicador.

Foram identificadas 53 bases de dados nos 25 artigos selecionados. Elas foram separadas em 9 categorias e estão relacionadas no Quadro 6. A base de dados abertos mais utilizada nos artigos foi o OpenStreetMap, servindo como fonte para os dados em 17 dos 25 artigos. O OpenStreetMap é uma plataforma mantida por uma comunidade voluntária que mantém atualizados os dados referentes ao mapeamento das cidades. Outra base de dados que se destaca é o Google Earth Engine, sendo utilizada em 7 artigos. Esta base de dados mantida pelo Google disponibiliza Imagens de Satélite, combinando-as com dados geoespaciais e recursos de análise em escala global. A plataforma WorldPop é a terceira mais utilizada nos artigos selecionados, aparecendo em 5 estudos. Esta é uma plataforma mantida pela Universidade de Southampton que mapeia populações em todo o mundo, fornecendo dados abertos de alta resolução sobre distribuições populacionais. Também se destacam os portais “tradicionais” de dados abertos governamentais, como aqueles que disponibilizam dados censitários de cada país.

Quadro 6: Bases de dados abertos identificadas nos artigos selecionados.

Tipos de dados	Base de dados	Detalhes	Abrangência	Artigos
Imagens de Satélite	Google Earth	-	Global	Zhong et al. (2020)
	Google Earth Engine	Sentinel-1 SAR GRD - Synthetic Aperture Radar Ground Range Detected	Global	Chen et al. (2021)
		Sentinel-2	Global	Chen et al. (2021); Hua et al. (2022)
		Suomi NPP VIIRS Day-Night Band imagery - Imagens noturnas	Global	Chen et al. (2021)
	USGS	Landsat 8 Satellite Image	Global	Wang, Han e Vries (2020); Mahabir et al. (2020)
		National Centers for Environmental Information	Suomi NPP VIIRS Day-Night Band imagery - Imagens noturnas	Global
	Global Stereo	WorldView2 satellite	Global	Grêt-Regamey et al. (2020)
		QuickBird	Global	Grêt-Regamey et al. (2020)
	Brazil Data Cube	Sentinel-2	Brasil	Adorno, Körting e Amaral (2023)
	DIDGI-INPE	CBERS-4A satellite - Panchromatic Wide-Scan Camera, Multispectral Camera, Wide Field Imager	Brasil	Adorno, Körting e Amaral (2023)
	National agriculture imagery program (NAIP)	Imagens de satélite agrícolas	EUA	Chen et al. (2021)

	Resource and Environmental Science Data Platform - RESDC	Landsat TM	China	Zhang, A. et al. (2021)
	National Elevation Dataset (NED)	Modelo digital de elevação	EUA	Chen et al. (2021)
	ASTER GDEM / NASA	Modelo digital de elevação	Global	Zhang, A. et al. (2021)
	USGS	Modelo digital de elevação - Shuttle Radar Topography Mission (SRTM)	EUA	Johnson et al. (2021)
Dados derivados de imagens de satélite	Global Human Settlements Layer (GHSL)	Dados de população	Global	Boeing et al. (2022); Liu et al. (2021); Carranza et al. (2022)
	Star Cloud Data Service Platform	Global artificial impervious area (GAIA) - área construída	Global	Johnson et al. (2021)
	gee-community-catalog	GFPLAIN250m - mapeamento de cheias	Global	Johnson et al. (2021)
	ESA WorldCover 10 m	baseado no Sentinel-2 - uso do solo	Global	Wu et al. (2023)
	TU Delft	Modelo digital de elevação - Current Height (AHN2)	Holanda	Wang, Han e Vries (2020)
Dados vetoriais espaciais	Open Street Map	Mapeamento colaborativo	Global	Boeing et al. (2022); Zhong et al. (2020); Chen et al. (2021); Biljecki e Chow (2022); Zhang, A. et al. (2021); Bittencourt et al. (2024); Blanco et al. (2024); Massano et al. (2023); Boeing (2021); Mahabir et al. (2020); Liu et al. (2021); Shartova, Mironova e Grischenko (2023); Carranza et al. (2022); He e He (2023); Zhou et al. (2024); Aparicio et al. (2024); Rui e Xu (2024)

	Baidu Maps	-	China	Zhang, A. et al. (2021); He e He (2023)
	Gaode Maps	-	China	Wu et al. (2023)
	Global Administrative Areas	Limites administrativos	Global	Biljecki e Chow (2022)
	GLOBIO	Global Roads Inventory Project (GRIP) - estradas	Global	Johnson et al. (2021)
Dados populacionais	WorldPop	-	Global	Chen et al. (2021); Biljecki e Chow (2022); Hua et al. (2022); Mahabir et al. (2020); Shartova, Mironova e Grischenko (2023)
	ORNL Population	LandScan Distribuição da população	Global	Zhang, A. et al. (2021); Mahabir et al. (2020)
Imagens de Ruas	Google Street View	-	Global	Hua et al. (2022)
	Baidu Street View	-	China	He e He (2023); Rui e Xu (2024)
Redes Sociais	Twitter	-	Global	Chen et al. (2021)
	Sina Weibo	-	China	He e He (2023)
	Flickr	-	Global	Mahabir et al. (2020)
	Foursquare	-	Global	Rosés, Kadar e Malleson (2021)
Outros tipos de dados	CityGML	idades em 3D	Global	Blanco et al. (2024)
	European Commission	INSPIRE Data Specification on Geographical Grid Systems - dados censitários agregados	Europa	Blanco et al. (2024)
	OpenMobilityData	General Transit Feed Specification (GTFS) - transporte público	Global	Boeing et al. (2022); Liu et al. (2021)
	Mapa de Valores del Suelo de América Latina	Dados de valor do solo - mapa colaborativo	América Latina	Carranza et al. (2022)
	Numbeo	Dados de tráfego; Dados de custo de vida	Global	Aparicio et al. (2024)
	Darksky API	Dados meteorológicos	Global	Rosés, Kadar e Malleson (2021)
	United Nations, Department of Economic and Social Affairs	Dados de crescimento da cidade	Global	Aparicio et al. (2024)

Datasets	Cityscapes	Imagens para treinamento de machine learning	-	He e He (2023)
	ADE20K	Imagens para treinamento de machine learning	-	He e He (2023); Rui e Xu (2024)
Bases de dados abertos governamentais	Shenzhen open data portal	Vários dados	Shenzen	He e He (2023); Rui e Xu (2024)
	NYC Open Data portal	Vários dados	Nova York	Rosés, Kadar e Malleson (2021)
	Planning Department de Hong Kong	Vários dados	Hong Kong	Hua et al. (2022)
	Caltrans PeMS	dados de tráfego	Califórnia	Zhou et al. (2024)
	National Household Travel Survey	dados de viagens	EUA	Zhou et al. (2024)
	Property24	Dados de imobiliárias	Quênia	Mahabir et al. (2020)
	MajiData	Dados de água e saneamento	Quênia	Mahabir et al. (2020)
	Korea Accreditation Board (KAB)	Consumo de energia	Coréia do Sul	Choi e Yoon (2020)
	Dianping	Dados de empresas chinesas	China	Zhang, A. et al. (2021)
	Royal Netherlands Meteorological Institute (KNMI)	Dados meteorológicos	Holanda	Wang, Han e Vries (2020)
	European Data	Research Movements in the Netherlands (Ovin) - dados de movimentos	Holanda	Wang, Han e Vries (2020)
		Basic registration Large-scale Topography (BGT) - mapa topográfico	Holanda	Wang, Han e Vries (2020)
	Central Bureau of Statistics (CBS)	Dados censitários	Holanda	Wang, Han e Vries (2020)
		Soil usage file (BBG) data - Dados de uso do solo	Holanda	Wang, Han e Vries (2020)
	United States Census Bureau	Dados censitários	EUA	Rosés, Kadar e Malleson (2021)
Census Bureau American Community Survey		Boeing (2021); Zhou et al. (2024)		
Kenya Open Data	Dados censitários	Quênia	Mahabir et al. (2020)	

Fonte: os autores.

Identificou-se que muitos dos dados oferecidos pelas bases de dados abertas são provenientes de Imagens de Satélite, tanto o dado “bruto” (a imagem de satélite) quanto com algum tipo de tratamento (como modelo digital de elevação, por exemplo). Além disso, os dados abertos governamentais, mesmo que muitas vezes atualizados com pouca frequência, ainda apresentam protagonismo nos estudos. No entanto, plataformas colaborativas como o OpenStreetMap e o Mapa de Valores del Suelo de América Latina também ganham destaque.

Neste sentido, observamos dois tipos de plataformas de dados abertos: 1) *plataformas top-down*, que disponibilizam dados coletados pelo poder governamental ou outras instituições; e 2) plataformas *bottom-up*, que são construídas coletivamente, por meio da colaboração dos usuários.

Também foram identificadas diversas ferramentas e técnicas utilizadas nos artigos, que podem ser encontradas no Quadro 7. Nesta relação, foram consideradas apenas as ferramentas e técnicas que foram citadas nos artigos analisados. Em relação ao processamento dos dados, ferramentas já estabelecidas há bastante tempo na área de dados para Planejamento Urbano destacam-se, como as ferramentas GIS, utilizadas em 11 dos 25 artigos. Porém, linguagens de programação como Python (utilizada em 11 artigos) e R (utilizada em 4 artigos) também se encontram em evidência. Além disso, técnicas de estatística também foram muito utilizadas, em especial o Índice de Moran, técnica para medir autocorrelação espacial, empregue em 6 estudos. Técnicas de Machine Learning e Deep Learning também mostram relevância, em particular a técnica Random Forest, para combinação de árvores de decisão, e a técnica k-means para clusterização.

Quadro 7: ferramentas e técnicas utilizadas nos artigos.

Linguagens de programação				
Grupo	Artigos	Bibliotecas / Pacotes	Função	Artigos
Python		GeoPandas	manipulação de dados espaciais	Blanco et al. (2024); Liu et al. (2021)
	Chen et al. (2021); Biljecki e Chow (2022);	scikit-learn	machine learning	Blanco et al. (2024); Mahabir et al. (2020); Rosés, Kadar e Malleson (2021); Rui e Xu (2024)
	Bittencourt et al. (2024); Blanco et al. (2024);	OSMnx	download de dados do OSM	Massano et al. (2023); Liu et al. (2021)
	Massano et al. (2023); Boeing et al. (2021); Mahabir et al. (2020); Liu et al. (2021);	pandana	medir acessibilidade a viagens	Liu et al. (2021)
	Rosés, Kadar e Malleson (2021); He e He (2023); Rui e Xu (2024)	pandas	análise e manipulação de dados	Liu et al. (2021)
		pysal	análise espacial	Liu et al. (2021)
		scipy	computação científica	Liu et al. (2021)
		numpy	funções matemáticas	Liu et al. (2021)
		NetworkX	manipulação de redes complexas	Liu et al. (2021); Rui e Xu (2024)
		mesa	cria agent-based models	Rosés, Kadar e Malleson (2021)
	statsmodels	estatísticas	Rosés, Kadar e Malleson (2021)	
	WeiboSpider	usar dados da rede social	He e He (2023)	
R	Biljecki e Chow (2022); Mahabir et al. (2020);	SITS – R	processamento de Satellite Image Time-Series	Adorno, Körting e Amaral (2023)

	Adorno, Körting e Amaral (2023); Carranza et al. (2022)	osmdata rgee sf	importar dados OSM importar dados google earth engine codificar dados vetoriais espaciais	Carranza et al. (2022) Carranza et al. (2022) Carranza et al. (2022)
Ambientes de processamento				
Grupo	Artigos	Plugins	Função	Artigos
GIS (ArcGIS, QGIS, TerraView)	Biljecki e Chow (2022); Wang, Han e Vries (2020); Hua et al. (2022); Grêt-Regamey et al. (2020); Blanco et al. (2024);	sDNA (Hillier and Hanson)	sintaxe espacial	Rui e Xu (2024)
	Massano et al. (2023); Wu et al. (2023); Adorno, Körting e Amaral (2023);	GeoDMA	mineração de dados de sensoriamento remoto	Adorno, Körting e Amaral (2023)
	Shartova, Mironova e Grischenko (2023); He e He (2023); Rui e Xu (2024)	HCMGIS-plugin	importar dados do OSM	Blanco et al. (2024)
		FlexiGIS	análise de sistemas energéticos	Blanco et al. (2024)
		Orfeo ToolBox	processamento de imagens de sensoriamento remoto	Adorno, Körting e Amaral (2023)
PostgreSQL/PostGIS		SQL	banco de dados	Biljecki e Chow (2022); Liu et al. (2021)
Outros Softwares		Flexible Modelling Framework	microsimulação espacial	Wang, Han e Vries (2020)
		MATLAB	cálculo numérico	Hua et al. (2022)
		Stata	análise estatística	Hua et al. (2022)
		ST-RISK	análise de riscos de terremoto	Bittencourt et al. (2024)
		MGWR 2.2	calcular GWR	Wu et al. (2023)
		GeoDa	estatísticas	Wu et al. (2023)
		WEKA	estatísticas	Johnson et al. (2021)
		MATSim	ABM	Zhou et al. (2024)
Técnicas				
Grupo	Especificação	Função	Artigos	
Estatísticas		Regressão logística	probabilidade de ocorrência de um evento	Wang, Han e Vries (2020); Mahabir et al. (2020); Johnson et al. (2021)
		geographically weighted regression (GWR)	análise de relação espacial / regressão linear	Zhang, A. et al. (2021); Wu et al. (2023)

	ordinary least squares (OLS)	regressão linear	Zhang, A. et al. (2021); Wu et al. (2023)
	índice de Moran	medir autocorrelação espacial	Zhang, A. et al. (2021); Wu et al. (2023); Shartova, Mironova e Grischenko (2023); Carranza et al. (2022); He e He (2023); Rui e Xu (2024)
	Hot Spot Analysis Getis-Ord Gi*	identifica pontos frios e quentes espacialmente	Grêt-Regamey et al. (2020)
	Spearman' Rho	descreve a relação entre duas variáveis	Grêt-Regamey et al. (2020)
	Discriminant analysis	classificar dados em grupos	Mahabir et al. (2020)
	Pearson correlation	mede a relação entre duas variáveis	Wu et al. (2023); Rui e Xu (2024)
	LISA	indicador local de autocorrelação espacial	Shartova, Mironova e Grischenko (2023)
	IPF method	alinhar distribuição de dados	Zhou et al. (2024)
	Multiscale Geographically Weighted Regression (MGWR)	regressão linear	Rui e Xu (2024)
	Simpson index	medir diversidade de um local	Rui e Xu (2024)
Machine Learning	Random Forest	combinação de árvores de decisão	Chen et al. (2021); Hua et al. (2022); Blanco et al. (2024); Adorno, Körting e Amaral (2023); Carranza et al. (2022); Rui e Xu (2024)
	Extremely Randomized Tree	cria árvores de decisão aleatórias	Chen et al. (2021)
	LightGBM	estrutura de aumento de gradiente	Chen et al. (2021), Rui e Xu (2024)
	CatBoost	estrutura de aumento de gradiente	Chen et al. (2021)
	Redes Neurais		Chen et al. (2021)
	Algoritmo K-means	clusterização	Zhang, A. et al. (2021); Choi e Yoon (2020); Bittencourt et al. (2024); Rui e Xu (2024)
	k-medoids	clusterização	Grêt-Regamey et al. (2020)
	Quantile Random Forest (QRF)	arvores de decisão	Carranza et al. (2022)

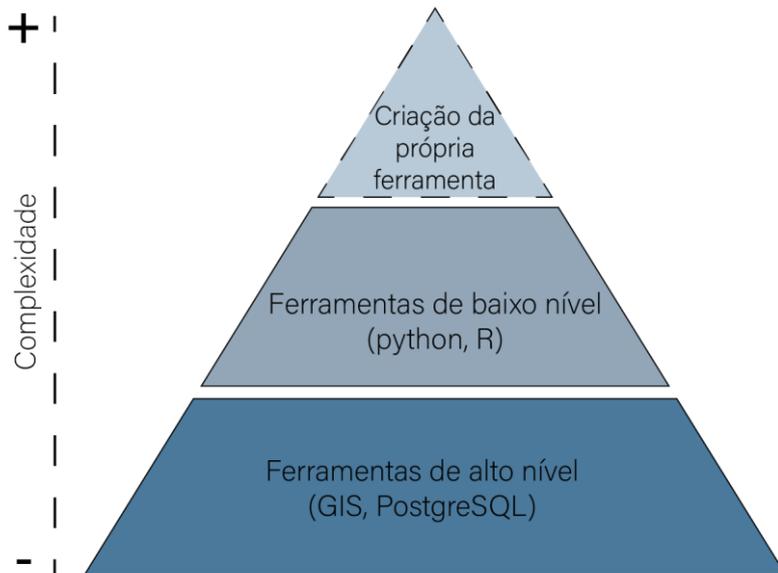
	Gradient Boosting Model (GBM)	classificação de árvores de decisão	Carranza et al. (2022)
	Video Propagation and Label Relaxation (VPLR)	segmentação semântica de imagens	He e He (2023)
	Fast Segmentation Convolutional Neural Network (Fast-SCNN)	segmentação semântica de imagens	He e He (2023)
	DeepLabV3	segmentação semântica de imagens	He e He (2023), Rui e Xu (2024)
	ResNeSt	segmentação semântica de imagens	He e He (2023)
	XGBoost	classificação e árvores de regressão	Rui e Xu (2024)
	K-nearest neighbors (KNN)	cálculo de distância entre grafos	Rui e Xu (2024)
	back propagation	redes neurais	Rui e Xu (2024)
	suppor vector machines (SVMs)	classificação e regressão	Rui e Xu (2024)
Deep Learning	convolutional neural networks (CNN)	análise de imagens	Zhong et al. (2020)
	Modelo Word2Vec	similaridade de texto	Zhong et al. (2020)
	bag-of-words (BoW) model	extrair recursos de texto	Zhong et al. (2020)
	treinamento para reconhecimento de imagens	segmentação semântica de imagens	Hua et al. (2022)
Data Mining	See5	árvore de decisão	Mahabir et al. (2020)
Algoritmos	Dijkstra algorithm	caminho mais curto em grafos	Liu et al. (2021); Rosés, Kadar e Malleson (2021)
	DBSCAN	clusterização	Wu et al. (2023)
Modelos computacionais ou paramétricos	cellular automata		Johnson et al. (2021)
	Agent-based models (ABMs)	simulação de agentes	Rosés, Kadar e Malleson (2021); Zhou et al. (2024)
	C-Fix Model	estimar capacidade de absorção de carbono	Wang, Han e Vries (2020)

Fonte: os autores.

Através da análise das ferramentas utilizadas nas pesquisas dos artigos selecionados, verificamos que há várias formas de processar os mesmos tipos de dados. Assim, identificamos dois tipos de ferramentas: 1) Ferramentas de Alto Nível: são ferramentas que são mais abstraídas, ou seja, há uma redução da complexidade para facilitar o entendimento. Estes recursos normalmente são menos customizáveis, pois já possuem uma gama de funcionalidades padrão. Neste grupo, encontram-se os “Ambientes de processamento” do Quadro 9. 2) Ferramentas de Baixo Nível: são aquelas menos abstraídas, que necessitam

conhecimentos mais específicos para sua utilização. Estas ferramentas são mais customizáveis, e permitem o processamento dos mesmos dados do grupo de ferramentas anterior, porém de forma mais complexa, como as linguagens de programação. Neste sentido, pode-se mencionar ainda um terceiro nível de ferramentas (Figura 10), que são aquelas criadas pelos próprios autores. Entre os artigos selecionados, Boing et al. (2022) cria sua própria ferramenta.

Figura 10: nível de complexidade das ferramentas.



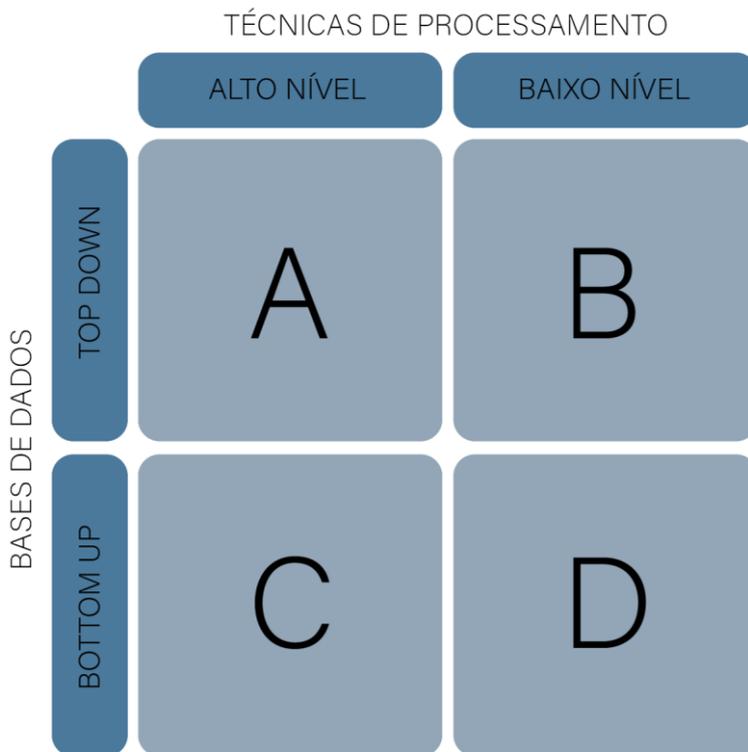
Fonte: os autores.

Assim, considerando os tipos de bases de dados identificadas e as ferramentas de processamento, identificamos quatro cenários relativos o uso de dados abertos para o estudo das cidades (Figura 11):

- Cenário A) Neste cenário, utiliza-se ferramentas de alto nível (menos complexas) para manipulação de dados coletados de maneira *top-down*. Este cenário é o mais tradicional para Planejamento Urbano, já que utiliza ferramentas como o GIS para o tratamento de dados de satélite ou governamentais.
- Cenário B) Uso de ferramentas de baixo nível para manipulação de dados *top-down*. Neste cenário, as ferramentas de processamento são mais customizáveis (porém mais complexas), permitindo análises diferentes para dados tradicionais.
- Cenário C) Uso de ferramentas de alto nível para manipulação de dados coletados de maneira *bottom-up*. Neste caso, o “afastamento” do modo tradicional de análise de dados se dá pela fonte dos dados, que é construída coletivamente.

- Cenário D) Uso de ferramentas de alto nível para manipulação de dados *bottom-up*. Este cenário, na perspectiva do Planejamento Urbano, é o mais inovador, pois utiliza ferramentas e dados não tradicionais.

Figura 11: Cenários possíveis para o uso de dados para estudo das cidades.



Fonte: os autores.

Por utilizarem dados abertos, a maioria dos estudos mostra uma alta transferibilidade. Inclusive, diversos trabalhos (Boeing et al., 2022; Zhong et al., 2020; Biljecki e Chow, 2022; Liu et al., 2021; Zhou et al., 2024) apresentam frameworks de código aberto para a replicação do estudo. Neste sentido, considerou-se que um estudo tem alta transferibilidade quando todas as bases de dados utilizadas são abertas e de abrangência global, ou utilizam dados “tradicionais” do censo nacional, como população ou densidade demográfica, que normalmente estão disponíveis para todos os países. Dessa forma, alguns estudos (Wang, Han e Vries, 2020; Choi e Yoon, 2020; Rosés, Kadar e Malleson, 2021; He e He, 2023; Zhou et al., 2024; Rui e Xu, 2024) foram considerados de baixa transferibilidade por utilizarem dados que são abertos, porém muito específicos daquele país ou cidade, e que não estão disponíveis para todos os lugares.

CONCLUSÃO

Este artigo abordou como dados abertos podem ser utilizados para análise, planejamento, gestão e desenho urbano. Para atingir os objetivos da pesquisa, foi utilizado um método misto de investigação, considerando aspectos quantitativos e qualitativos. Primeiramente, foi

realizada análise bibliométrica para a avaliação externa dos artigos selecionados, com o objetivo de compreensão do campo de pesquisa. Esta análise bibliométrica considerou aspectos sobre as revistas, autores, países, palavras-chave e seu relacionamento. Em seguida, a pesquisa seguiu através de revisão de literatura, fazendo a análise interna de 25 artigos. A revisão de literatura foi do tipo integrativa e considerou as características dos artigos, a escala de análise, abordagem do estudo, os conceitos relacionados ao campo de Planejamento Urbano que foram utilizados, as bases de dados abertas, as ferramentas utilizadas e o nível de transferibilidade do estudo.

O mapeamento do campo de pesquisa, feito através da análise bibliométrica, indica que o campo é multidisciplinar, multitemático e encontra-se em formação. No entanto, apresenta um crescimento nos últimos cinco anos, indicando o aumento da viabilidade da análise urbana baseada em dados abertos. Desta forma, considerando o constante desenvolvimento de novas tecnologias e a crescente disponibilidade de dados, pode-se considerar que há uma tendência de crescimento do campo. Em relação à análise interna dos artigos, os estudos possuem caráter empírico e utilizam conceitos do campo de planejamento urbano para definição das análises que serão apoiadas pelos dados abertos. Muitos estudos também se beneficiam da disponibilidade de dados para aplicar os mesmos métodos em cidades diferentes, comparando os resultados e permitindo a validação do método.

Além disso, foram identificados dois tipos de bases de dados abertas, relativos ao método de coleta dos dados: *top-down* e *bottom-up*. Entre as ferramentas utilizadas nos artigos, também podemos distingui-las em dois tipos: ferramentas de alto nível (menos complexas) e ferramentas de baixo nível (mais complexas). A partir destas informações, criamos uma matriz com quatro cenários possíveis, combinando as bases de dados e as ferramentas disponíveis. A partir disso, considera-se que as bases de dados abertas podem fornecer recursos para o monitoramento urbano, já que fornecem séries temporais de dados. Portanto, pode-se concluir que este artigo forneceu subsídios para o monitoramento urbano através de dados abertos.

FINANCIAMENTO

Este trabalho foi financiado pela Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina (FAPESC), Edital 19/2024.

REFERÊNCIAS

ADORNO, Bruno Vargas; KÖRTING, Thales Sehn; AMARAL, Silvana. Contribution of time-series data cubes to classify urban vegetation types by remote sensing. **Urban Forestry & Urban Greening**, [S.L.], v. 79, p. 127817, jan. 2023.

APARICIO, Joao Tiago; ARSENIO, Elisabete; SANTOS, Francisco C.; HENRIQUES, Rui. Walkability defined neighborhoods for sustainable cities. **Cities**, [S.L.], v. 149, p. 104944, jun. 2024.

BATTY, Michael. **The New Science of the Cities**. Cambridge: The Mit Press, 2013.

BIBRI, Simon Elias. Data-driven smart sustainable cities of the future: an evidence synthesis approach to a comprehensive state-of-the-art literature review. **Sustainable Futures**, [S.L.], v. 3, p. 100047, 2021.

BILJECKI, Filip; CHOW, Yoong Shin. Global Building Morphology Indicators. **Computers, Environment And Urban Systems**, [S.L.], v. 95, p. 101809, jul. 2022.

BITTENCOURT, João Carlos N.; COSTA, Daniel G.; PORTUGAL, Paulo; VASQUES, Francisco. A data-driven clustering approach for assessing spatiotemporal vulnerability to urban emergencies. **Sustainable Cities And Society**, [S.L.], v. 108, p. 105477, ago. 2024.

BLANCO, Luis; ALHAMWI, Alaa; SCHIRICKE, Björn; HOFFSCHMIDT, Bernhard. Data-driven classification of Urban Energy Units for district-level heating and electricity demand analysis. **Sustainable Cities And Society**, [S.L.], v. 101, p. 105075, fev. 2024.

BOEING, Geoff et al. Using open data and open-source software to develop spatial indicators of urban design and transport features for achieving healthy and sustainable cities. **The Lancet Global Health**, [S.L.], v. 10, n. 6, p. 907-918, jun. 2022.

BOEING, Geoff. Off the Grid...and Back Again? **Journal Of The American Planning Association**, [S.L.], v. 87, n. 1, p. 123-137, 19 out. 2021.

BURNS, Ryan; ANDRUCKI, Max. Smart cities: who cares?. **Environment And Planning A: Economy and Space**, [S.L.], v. 53, n. 1, p. 12-30, 14 jul. 2020.

CALLON, Michel; COURTIAL, Jean-Pierre; TURNER, William A.; BAUIN, Serge. From translations to problematic networks: an introduction to co-word analysis. **Social Science Information**, [S.L.], v. 22, n. 2, p. 191-235, mar. 1983.

CALLON, Michel; LAW, John; RIP, Arie. Qualitative Scientometrics. In: CALLON, Michel; LAW, John; RIP, Arie. **Mapping the dynamics of science and technology**. London: Macmillan Press Ltd, 1986. p. 103-123.

CARRANZA, Juan Pablo; PIUMETTO, Mario Andrés; LUCCA, Carlos María; SILVA, Everton da. Mass appraisal as affordable public policy: open data and machine learning for mapping urban land values. **Land Use Policy**, [S.L.], v. 119, p. 106211, ago. 2022.

CHAKRABORTY, Arnab; WILSON, Bev; SARRAF, Saket; JANA, Arnab. Open data for informal settlements: toward a user's guide for urban managers and planners. **Journal Of Urban Management**, [S.L.], v. 4, n. 2, p. 74-91, dez. 2015.

CHEN, Bin; TU, Ying; SONG, Yimeng; THEOBALD, David M.; ZHANG, Tao; REN, Zhehao; LI, Xuecao; YANG, Jun; WANG, Jie; WANG, Xi; GONG, Peng; BAI, Yuqi; BING, Xu. Mapping essential urban land use categories with open big data: results for five metropolitan areas in the united states of america. **Isprs Journal Of Photogrammetry And Remote Sensing**, [S.L.], v. 178, p. 203-218, ago. 2021.

CHOI, Sebin; YOON, Sungmin. Energy signature-based clustering using open data for urban building energy analysis toward carbon neutrality: a case study on electricity change under covid-19. **Sustainable Cities And Society**, [S.L.], v. 92, p. 104471, maio 2023.

DE BELLIS, Nicola. **Bibliometrics and citation analysis: from the science citation index to cybermetrics**. Plymouth (UK): Scarecrow Press, 2009.

GRÊT-REGAMEY, Adrienne; GALLEGUILLOS-TORRES, Marcelo; DISSEGNA, Angela; WEIBEL, Bettina. How urban densification influences ecosystem services—a comparison between a temperate and a tropical city. **Environmental Research Letters**, [S.L.], v. 15, n. 7, p. 075001, 18 jun. 2020.

GUERRI, Giulia; CRISCI, Alfonso; MESSERI, Alessandro; CONGEDO, Luca; MUNAFÒ, Michele; MORABITO, Marco. Thermal Summer Diurnal Hot-Spot Analysis: the role of local urban features layers. **Remote Sensing**, [S.L.], v. 13, n. 3, p. 538, 2 fev.

HE, Xuan; HE, Sylvia Y.. Using open data and deep learning to explore walkability in Shenzhen, China. **Transportation Research Part D: Transport and Environment**, [S.L.], v. 118, p. 103696, maio 2023.

HUA, Junyi; CAI, Meng; SHI, Yuan; REN, Chao; XIE, Jing; CHUNG, Lamuel Chi Hay; LU, Yi; CHEN, Long; YU, Zhaowu; WEBSTER, Chris. Investigating pedestrian-level greenery in urban forms in a high-density city for urban planning. **Sustainable Cities And Society**, [S.L.], v. 80, p. 103755, maio 2022.

JAIN, Garima; ESPEY, Jessica. Lessons from nine urban areas using data to drive local sustainable development. **npj Urban Sustainability**, v. 2, n. 7, 2022.

JOHNSON, Brian A.; ESTOQUE, Ronald C.; LI, Xuecao; KUMAR, Pankaj; DASGUPTA, Rajarshi; AVTAR, Ram; MAGCALE-MACANDOG, Damasa B.. High-resolution urban change modeling and flood exposure estimation at a national scale using open geospatial data: a case study of the philippines. **Computers, Environment And Urban Systems**, [S.L.], v. 90, p. 101704, nov. 2021.

KANDT, Jens; BATTY, Michael. Smart cities, big data and urban policy: towards urban analytics for the long run. **Cities**, [S.L.], v. 109, p. 102992, fev. 2021.

KESSLER, M. M.. Bibliographic coupling between scientific papers. **American Documentation**, [S.L.], v. 14, n. 1, p. 10-25, jan. 1963.

KITCHIN, Rob; MOORE-CHERRY, Niamh. Fragmented governance, the urban data ecosystem and smart city-regions: the case of metropolitan boston. **Regional Studies**, [S.L.], v. 55, n. 12, p. 1913-1923, 30 mar. 2020.

KOURTIT, Karima; NIJKAMP, Peter; STEENBRUGGEN, John. The significance of digital data systems for smart city policy. **Socio-Economic Planning Sciences**, [S.L.], v. 58, p. 13-21, jun. 2017.

LI, Han; CHEN, Peijun; GRANT, Richard. Built environment, special economic zone, and housing prices in Shenzhen, China. **Applied Geography**, [S.L.], v. 129, p. 102429, abr. 2021.

LIU, Shiqin; HIGGS, Carl; ARUNDEL, Jonathan; BOEING, Geoff; CERDERA, Nicholas; MOCTEZUMA, David; CERIN, Ester; ADLAKHA, Deepti; LOWE, Melanie; GILES-CORTI, Billie. A Generalized Framework for Measuring Pedestrian Accessibility around the World Using Open Data. **Geographical Analysis**, [S.L.], v. 54, n. 3, p. 559-582, 19 maio 2021.

LEI, Binyu; STOUFFS, Rudi; BILJECKI, Filip. Assessing and benchmarking 3D city models. **International Journal Of Geographical Information Science**, [S.L.], v. 37, n. 4, p. 788-809, 8 nov. 2022.

LUDWIG, Leandro. **A construção cognitiva do planejamento urbano a partir da ideia de matrizes de referência**: os conflitos normativos, políticos, ambientais e tecnológicos do espaço urbano. 2022. 261 f., il. Tese (Doutorado em Desenvolvimento Regional) - Programa de Pós-Graduação em Desenvolvimento Regional, Universidade Regional de Blumenau, Blumenau, 2022.

MAHABIR, Ron; AGOURIS, Peggy; STEFANIDIS, Anthony; CROITORU, Arie; CROOKS, Andrew T.. Detecting and mapping slums using open data: a case study in kenya. **International Journal Of Digital Earth**, [S.L.], v. 13, n. 6, p. 683-707, jun 2018.

MASSANO, Marco; MACII, Enrico; LANZINI, Andrea; PATTI, Edoardo; BOTTACCIOLI, Lorenzo. A GIS Open-Data Co-Simulation Platform for Photovoltaic Integration in Residential Urban Areas. **Engineering**, [S.L.], v. 26, p. 198-213, jul. 2023.

MATTEDI, Marcos; SPIESS, Maiko. **O pântano cognitivo**. Blumenau: Edifurb, 2023.

NOYONS, E.C.M.; MOED, H.F.; LUWEL, M.. Combining mapping and citation analysis for evaluative bibliometric purposes: a bibliometric study. **Journal Of The American Society For Information Science**, [S.L.], v. 50, n. 2, p. 115-131, 1999.

QIANG, Dan; ZHANG, Lingzhu; HUANG, Xiaotong. Quantitative Evaluation of TOD Performance Based on Multi-Source Data: a case study of shanghai. **Frontiers In Public Health**, [S.L.], v. 10, 21 fev. 2022.

ROEMER, Robin Chin; BORCHARDT, Rachel. **Meaningful Metrics**: a 21st-century librarian's guide to bibliometrics, altmetrics, and research impact. Chicago: Association Of College & University Libraries., 2015.

ROSÉS, Raquel; KADAR, Cristina; MALLESON, Nick. A data-driven agent-based simulation to predict crime patterns in an urban environment. **Computers, Environment And Urban Systems**, [S.L.], v. 89, p. 101660, set. 2021.

RUI, Jin; XU, Yuhan. Beyond built environment: unveiling the interplay of streetscape perceptions and cycling behavior. **Sustainable Cities And Society**, [S.L.], v. 109, p. 105525, ago. 2024.

SHARTOVA, N.V.; MIRONOVA, E.e.; GRISCHENKO, M. Yu.. Spatial disparities of street walkability in Moscow in the context of healthy urban environment. **Cities**, [S.L.], v. 141, p. 104469, out. 2023.

SMALL, Henry. Co-citation in the scientific literature: a new measure of the relationship between two documents. **Journal Of The American Society For Information Science**, [S.L.], v. 24, n. 4, p. 265-269, jul. 1973

SMALL, Henry. Visualizing science by citation mapping. **Journal Of The American Society For Information Science**, [S.L.], v. 50, n. 9, p. 799-813, 1999.

SNYDER, Hannah. Literature review as a research methodology: an overview and guidelines. **Journal Of Business Research**, [S.L.], v. 104, p. 333-339, nov. 2019.

STEIL, A. V. (2021). **Diferenças e similaridades entre revisão sistemática, revisão de escopo, revisão integrativa e revisão narrativa** [Material de apoio de aula online]. Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina.

TORRACO, Richard J.. Writing Integrative Literature Reviews: guidelines and examples. **Human Resource Development Review**, [S.L.], v. 4, n. 3, p. 356-367, set. 2005.

WANG, Gengzhe; HAN, Qi; VRIES, Bauke de. A geographic carbon emission estimating framework on the city scale. **Journal Of Cleaner Production**, [S.L.], v. 244, p. 118793, jan. 2020.

WU, Zihao; TONG, Ziyu; WANG, Mingzhu; LONG, Qianhui. Assessing the impact of urban morphological parameters on land surface temperature in the heat aggregation areas with spatial heterogeneity: a case study of nanjing. **Building And Environment**, [S.L.], v. 235, p. 110232, maio 2023.

ZHANG, Anqi; LI, Weifeng; WU, Jiayu; LIN, Jian; CHU, Jianqun; XIA, Chang. How can the urban landscape affect urban vitality at the street block level? A case study of 15 metropolises in China. **Environment And Planning B: Urban Analytics and City Science**, [S.L.], v. 48, n. 5, p. 1245-1262, 18 maio 2021.

ZHANG, Peng; HU, Shougeng; LI, Weidong; ZHANG, Chuanrong; YANG, Shengfu; QU, Shijin. Modeling fine-scale residential land price distribution: an experimental study using open data and machine learning. **Applied Geography**, [S.L.], v. 129, p. 102442, abr. 2021.

ZHONG, Yanfei; SU, Yu; WU, Siqi; ZHENG, Zhendong; ZHAO, Ji; MA, Ailong; ZHU, Qiqi; YE, Richen; LI, Xiaoman; PELLIKKA, Petri; ZHANG, Liangpei. Open-source data-driven urban land-use mapping integrating point-line-polygon semantic objects: a case study of chinese cities. **Remote Sensing Of Environment**, [S.L.], v. 247, p. 111838, set. 2020.

ZHOU, Yirong; LIU, Xiaoyue Cathy; CHEN, Bingkun; GRUBESIC, Tony; WEI, Ran; WALLACE, Danielle. A data-driven framework for agent-based modeling of vehicular travel using publicly available data. *Computers, Environment And Urban Systems*, [S.L.], v. 110, p. 102095, jun. 2024.

ZUPIC, Ivan; ČATER, Tomáš. Bibliometric Methods in Management and Organization. **Organizational Research Methods**, [S.L.], v. 18, n. 3, p. 429-472, 22 dez. 2014.