# **RESEARCH ARTICLE**

# **Open Access**



# Latent Dirichlet Allocation (LDA) topic models for Space Syntax studies on spatial experience

Ju Hyun Lee<sup>1\*</sup> and Michael J. Ostwald<sup>1</sup>

# Abstract

Spatial experience has been extensively researched in various fields, with Space Syntax being one of the most widely used methodologies. Multiple Space Syntax techniques have been developed and used to quantitively examine the relationship between spatial configuration and human experience. However, due to the heterogeneity of syntactic measures and experiential issues in the built environment, a systematic review of socio-spatial topics has yet to be developed for Space Syntax research. In response to this knowledge gap, this article employs an 'intelligent' method to classify and systematically review topics in Space Syntax studies on spatial experience. Specifically, after identifying 66 articles using the 'Preferred Reporting Items for Systematic reviews and Meta-Analyses' (PRISMA) framework, this research develops generative probabilistic topic models to classify the articles using the Latent Dirichlet Allocation (LDA) method. As a result, this research automatically generates three architectural topics from the collected literature data (A1. Wayfinding behaviour, A2. Interactive accessibility, and A3. Healthcare design) and three urban topics (U1. Pedestrian movement, U2. Park accessibility, and U3. Cognitive city). Thereafter it gualitatively examines the implications of the data and its LDA classification. This article concludes with an examination of the limitations of both the methods and the results. Along with demonstrating a methodological innovation (combining PRISMA with LDA), this research identifies critical socio-spatial concepts and examines the complexity of Space Syntax applications. In this way, this research contributes to future Space Syntax research that empirically investigates the relationships between syntactic and experiential variables in architectural and urban spaces. The findings support a detailed discussion about research gaps in the literature and future research directions.

Keywords Topic modelling, Systematic review, Space Syntax, Latent Dirichlet Allocation (LDA), PRISMA

# Introduction

Traditional social-spatial theories—for example, Gibson's ground theory (1950) and Kaplan and Kaplan's information theory (1989)—typically used empirical methods to examine the collective characteristics of human responses to architectural and urban space. Such psychological methods have been shown to capture an environment's social and behavioural properties. Still, their results are not always reproducible or able to be used for predictive modelling (Stamps 2005). In contrast, the Space Syntax set of methods, which use topology and graph theory mathematics, is a quantitative method for understanding and modelling the relationships between space and social patterns. Because Space Syntax approaches are independent of social parameters (Vaughan 2015), they can overcome some of the limitations of traditional social-spatial research, which relies more on human data than spatial data. As a result of its emphasis on reproducibility, modelling and spatial data, Space Syntax theories and methods have been widely



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

<sup>\*</sup>Correspondence:

Ju Hyun Lee

juhyun.lee@unsw.edu.au

<sup>&</sup>lt;sup>1</sup> School of Built Environment, Faculty of Arts, Design and Architecture,

The University of New South Wales, Sydney, NSW 2052, Australia

used to explore complex spatial experiences in the built environment.

Since the earliest syntactic spatial theories (Hanson 1998; Hillier 1996, 1999; Hillier and Hanson 1984), researchers have demonstrated that spatial configurations have social consequences (Hillier and Hanson 1984). Hundreds of studies have been completed using Space Syntax methods that examine socio-spatial relationships in cities, urban spaces, and buildings. Furthermore, recent empirical studies have measured various experiential or behavioural data to statistically examine syntactic and experiential relationships in architectural and urban spaces (Lee et al. 2023). Research topics in these experiential studies are, however, sufficiently diverse and varied that they have largely resisted attempts to synthesise knowledge from them. The present study addresses this knowledge gap, the lack of a collective view of syntactic and experiential issues, which is significant for current architectural and urban research and the development of future socio-spatial studies. An important first step to addressing this research problem is to answer the question: "Which topics have been investigated in Space Syntax research on spatial experience?" To address this question, the present research statistically identifies important syntactic and experiential topics in Space Syntax research and then systematically reviews these to determine their current and future implications.

Achieving this objective requires identifying and classifying socio-spatial topics investigated in the past before extrapolating their implications for future research. The present research starts by collating relevant Space Syntax studies using the 'Preferred Reporting Items for Systematic reviews and Meta-Analyses' (PRISMA) framework (Moher et al. 2009; Page et al. 2021) and then it characterises the topics contained in this research using Latent Dirichlet Allocation (LDA) (Asmussen and Møller 2019; Blei et al. 2003; Kee et al. 2019). LDA, an automatic method used for topic modelling, is derived from advanced natural language processing (NLP) models. The generative process of LDA has several significant advantages over most current methods, including a capacity to reveal hidden topics and new insights (Jelodar et al. 2019), which is why the present research uses it. Since LDA automatically generates a set of topics from the data, it is more efficient and arguably more rigorous than most common, subjective manual identification methods (Kee et al. 2019).

Although the PRISMA framework has been widely used in various research communities for systematic reviews, there are only a few PRISMA-based reviews of Space Syntax research (Sharmin and Kamruzzaman 2018). Furthermore, most systematic reviews of this type in architectural and urban sciences are regarded as "scoping" rather than "synthesising" studies (Lee et al. 2023), in part because they lack a rigorous process for

2023), in part because they lack a rigorous process for extracting topics or themes from the set of works. This is why the present article employs a novel method, wherein two frameworks—PRISMA for scoping works and LDA for automatic, computational topic modelling—are combined to both "scope" and begin to "synthesise" Space Syntax research about spatial experience.

The following section presents a detailed background to Space Syntax, its theory, and its techniques. Thereafter, a methodology section illustrates the research framework consisting of data collection and topic modelling. This section includes a detailed description of LDA's origins and specific application. The following sections report the results of LDA analysis, focusing on two subsets of the data, architectural and urban topics, and their implications. Three architectural topics are identified using LDA, and the research within these is critically discussed, identifying synergies, lessons, and knowledge gaps. Then three urban spatial topics are reviewed, drawing out deeper observations about the relationships between them, which LDA assists in understanding. This article concludes with a discussion of syntactic and experiential topics for future research.

#### Space syntax theory and techniques

Focusing on two-dimensional (2D) spatial logic, a range of Space Syntax techniques have been developed and applied to quantify the properties of architectural and urban spaces. Three of the fundamental socio-spatial principles that underly Space Syntax theory are that 'people move in lines', 'interact in convex space', and 'see changing visual fields' (Hillier and Vaughan 2007; Lee and Ostwald 2020), which are referred to as alpha analysis (or ALA), gamma analysis (or CSA), and isovist analysis (ISA), respectively (Hillier 1996, 1999; Hillier and Hanson 1984). Most Space Syntax techniques (Fig. 1), developed from these three, operate by converting the spatial properties of the built environment into a set of nodes and edges in a map or graph, which is then mathematically measured to explore socio-spatial relationships. In this way, regardless of whether the method is applied to buildings or cities (Hillier and Hanson 1984), it can be used to identify spatial topologies and social relations. However, due to the innate simplicity of topological abstraction, the theory has continued to evolve since first being proposed in the 1980s. As such, new Space Syntax techniques and approaches have been developed and tested by many researchers to address limitations of the theory and specific contentious issues such as the impacts of distance, height, scale and sinuosity on measurements (Netto 2016; Pafka et al. 2020; Ratti 2004).



Fig. 1 Space Syntax theory and related techniques

Probably, the most common syntactical method, ALA, examines 'straight' lines of human movement on streets or paths, which are treated as 'axial lines' on a 2D map or plan. In an axial map of a city, for example, street intersections become nodes and the streets are edges (or links) on an axial graph. The connectivity properties of these nodes and edges are then mathematically measured using graph theory (Lee and Ostwald 2020). In this way, ALA measurements consider the 'collective set of all the possible movements' in an environment (Hillier 1996) or the natural movement of 'very large collections of people' (Hillier 2012).

Complex human movements and spatial systems might not, however, be captured in this simplified topological view of the built environment. Consequently, the traditional ALA approach, often combined with geographical information system (GIS) data (Jiang and Liu 2009; Srinurak and Mishima 2017), has been expanded to improve its measures and their accuracies. For example, axial lines can be broken into axial segments and used to examine angular changes in direction (Turner 2007). This alternative to conventional ALA facilitates consideration of metric distance and navigational properties, a criticism of earlier versions of the approach (Abshirini and Koch 2016; di Bella et al. 2015). Furthermore, while distances can be topologically measured (Marcus and Koch 2016) and topological (geometric and angular) distances can even be superior to metric ones in some urban studies (Chiang and Li 2019; Hillier and Vaughan 2007), refinements of metric weighting (Kim and Piao 2017) and conventional metric-based measures (Hajrasouliha and Yin 2014; Scoppa and Peponis 2015) continue to be proposed. Interestingly, two syntactical measures-closeness centrality and betweenness centrality-are used interchangeably for 'integration' and 'choice', respectively, in many studies using ALA (di Bella et al. 2015; Jayasinghe et al. 2017; Jiang and Jia 2011; Morales et al. 2017; Omer and Goldblatt 2017; Omer and Jiang 2015). However, original centrality measures, developed for social network analysis (Feeman 1978), can be used to weight 'metric distance' (Lee and Ostwald 2020; Lee et al. 2022). Nonetheless, the syntactic versions of centralities are simply based on segmentation and radius (Pafka et al. 2020). Furthermore, street segments can be merged either via identification with the same name ('named street') or through a clear sense of continuity ('natural street') (Jiang and Claramunt 2004; Jiang and Liu 2009), with the latter potentially being superior for predicting human activities (Ma et al. 2018). Such street-based topological approaches are considered advanced versions of a conventional axial map.

In contrast, CSA and ISA-more commonly known as JPG and VGA, respectively-have been used for architectural analysis and have not had the same level of development as ALA. The JPG, a more diagrammatic variation of CSA, produces a 'planar graph' representation of 'plan morphology' (Steadman 1983), also known as a 'justified permeability map' (Hillier and Hanson 1984). The major syntactic idea underpinning CSA or JPG is that the enclosure of a space limits or shapes human behaviours (Hayward and Franklin 1974; Lee et al. 2016; Ostwald 2011). Thus, CSA/JPG has been used to examine the spatial configurations of plan layouts. The method does, however, usually produce a simple graph that is more limited in its capacity to measure experiential properties. For example, in the classical CSA or JPG, there cannot be two doors (connections or links) between the same two rooms (or nodes). A weighted and directed graph analysis can, however, overcome this limitation (Lee et al. 2022).

Interestingly, Hillier and his colleagues acknowledge this problem in a JPG, but not in their syntactic measures (Hillier et al. 1987). In contrast, both ISA and VGA are related through the use of an isovist—"the set of all points visible from a given vantage point in space" (Benedikt 1979, p. 47)—but they have different levels of empirical evidence and applications (Braaksma and Cook 1980; Turner et al. 2001). For example, VGA may provide better holistic measures like integration, while ISA seems to be more useful for capturing the spatial viability characteristics of specific isovists. Nonetheless, VGA is one of the most common Space Syntax methods and is often combined with ALA and/or agent-based simulation (ABS) to measure both visibility and permeability structures in the built environment.

Space Syntax derives a variety of quantitative descriptions of the topological properties of graphs or maps, largely using UCL Depthmap software (Turner 2001; Turner et al. 2001), which can then be correlated to various aspects of human behaviour. For example, integration is a normalised measure of distance from any space of origin to all others in a system (Hillier and Hanson 1984). Connectivity is the number of direct connections to other spaces or movement paths, and intelligibility is the correlation between connectivity and integration (Hölscher et al. 2012; Lee and Ostwald 2020). Choice is a global dynamic measure that reflects the "flow" or movement through space (Pagkratidou et al. 2020). In ALA and CSA, the syntactic properties of spatial configurations are usually calculated using topological values derived from axial and convex maps. In contrast, ISA can directly measure the properties of isovist viewsheds-e.g., isovist area, isovist perimeter, max radial, min radial, and Occlusivity-from architectural plans (Lee et al. 2017). VGA also identifies the syntactic properties such as integration and connectivity from visibility graphs or maps, in other words, visual integration and visual connectivity.

Because of the wide variety of recent syntactic measures developed for Space Syntax research (Sharmin and Kamruzzaman 2018), their interpretations can be more complicated than those of conventional topological analysis (Lee and Ostwald 2020). For example, the standard integration measure has generally been used to explore the spatial form of cities and their nonlocal properties (Hillier 1999). In contrast, Hillier and his colleagues (Hillier 2012; Hillier and Hanson 1984; Hillier and Iida 2005; Hillier et al. 2012) suggest that integration is more related to destination-based or 'to-movement' than choice or 'through-movement'. Thus, mathematically, both are sometimes interchangeably used for closeness centrality and betweenness centrality, respectively (Lee and Ostwald 2020), which can lead to confusion in the interpretation (Pafka et al. 2020). Conversely, these popular syntactic measures can be enhanced using segment analysis (SA) to consider metric ('least length'), geometric ('least angle change'), and topological ('fewest turns') properties as well (Hillier and Iida 2005). Evidence suggests that angular syntactic measures may be superior to metric or topological ones (Serra and Hillier 2019).

Space Syntax research is commonly used to compare 'syntactic' or spatial connectivity data, with human experiential or experimental data relating to movement, behaviour, and perception (Lee et al. 2023). Whilst human data can be collected by several research techniques-such as using questionnaires or interviewsvarious types of field observations are widely conducted to examine actual socio-spatial phenomena in the built environment. For example, movement patterns such as route and density are developed using 'gate counting' (Istiani et al. 2023; Khotbehsara et al. 2023; Mansouri and Ujang 2017; Omer and Goldblatt 2017; Shen and Wu 2022) and field observation (Can Traunmüller et al. 2023; Esposito et al. 2020; Koohsari et al. 2016; Liu et al. 2018; Nicoletta et al. 2022; Ozbil et al. 2021; Shatu et al. 2019; Yang et al. 2023; Zhai and Baran 2016), where ALA is dominant but VGA is also used as a syntactic method. Furthermore, GPS or electronic tracking (Domènech et al. 2020; Neo and Sagha-Zadeh 2017; Zhai et al. 2018; Zhang et al. 2020; Zheng et al. 2022) and video recording (Hölscher et al. 2012; Karthika et al. 2022; Tzeng and Huang 2009) are used for this purpose. In addition, snapshot observation is used to capture and categorise human behavioural data (e.g., playing, sitting, standing, talking, and walking) (Askarizad and Safari 2020; Can and Heath 2016üş and Yılmaz 2022; Zerouati and Bellal 2020). In this way, observation techniques are frequently used to develop human movement and behavioural data in Space Syntax research on spatial experience.

In summary, researchers argue that the adaption of syntactical measures to new aspects of socio-spatial theory should remain a priority (Netto 2016). Likewise, the socio-spatial dimension of human experience or behaviour in the built environment is an important area where complex new research is being undertaken, much of which is difficult to formalise or conceptualise in a collective way. For this reason, the classification of syntactic and experiential topics in Space Syntax research is the focus of the present research.

# **Research methodology and material** Research framework

As illustrated in Fig. 2, this study identifies relevant research articles using PRISMA and then classifies syntactic and experiential topics in the collected data using LDA topic modelling. Compared to the original four-phase data collection ('identification', 'screening',



Fig. 2 Research framework consisting of data collection and topic modelling

'eligibility', and 'included'), the PRISMA 2020 (Page et al. 2021) framework combines screening and eligibility into a single phase, which is used in the present data collection. Importantly, abstracts extracted from relevant research articles are used as data for topic modelling. Although different types of textual data could have an impact on topic modelling (Syed and Spruit 2017), abstract data is the most valuable and also the most frequently used in text-mining methods (Fang et al. 2018; Heimerl et al. 2014; Syed and Spruit 2017; Wu et al. 2012).

Whereas topic modelling in isolation collects text data or documents without any preliminary filtering (AlSumait et al. 2008; Blei et al. 2003; Kee et al. 2019), the present research collects document data more formally through PRISMA. After the preprocessing of the collected abstracts (tokenization, stopwords, and stemming), the topic modelling stage in Fig. 2 uses LDA, one of the most accepted topic modelling methods in a variety of fields (Fang et al. 2018; Wu et al. 2012). Finally, the present research infers and classifies syntactic and experiential topics. There are several possible open-source *Python* and *MATLAB* codes for this type of NLP and topic modelling. For the present application, after assessing several *Python* libraries, the final LDA algorithm was developed from Barber's work (Barber 2023).

#### Data collection

As this study is concerned with syntactic and experiential issues in Space Syntax research, the search string for online research databases used two initial keywords 'space syntax' and 'experience', to identify studies focussing on experiential data for correlational analysis. This study tested alternative keywords like 'behaviour' or 'social' to find appropriate socio-spatial issues, but their applications resulted in thousands of articles being identified, a structured review of which identified that many were irrelevant to the theme. Finally, the keywords 'experiment OR test' were included in the search string to target case studies rather than reviews or theoretical articles.

In November 2021, the first data collection, online database search, was conducted using nine online databases (Emerald Insight, JSTOR, MDPI, SAGE Journals, Science Direct, Taylor & Francis Online, UCL Discovery, Web of Science, and Wiley Online Library). These encapsulate primary refereed journals in architecture and urban design. The initial search identified 1317 articles. Following screening of the initial data by title and abstract, two researchers assessed the quality of articles for eligibility by an independent full-text review with a simple rating system. Finally, the online database search selected 26 articles. The second data collection, citation search, was then conducted to ensure all important papers were included in the identification process. From the 26 articles, the forward and backward citation searches identified 1801 potential records but finally included 12 articles. In this way, after a methodological pilot test with the initial data set, the third literature collection was done in June 2023, adding 28 articles. Collectively, this review examines 66 refereed journal articles (see also Additional file 1: Table S1).

The collected data consists of 27 architectural studies that deal with various architectural spaces, ranging from 'house' (Marquardt et al. 2011; Zeng et al. 2020) to 'hospital' (Alalouch and Aspinall 2007; Geng et al. 2021; Tzeng and Huang 2009), and 39 urban studies, the primary foci of which are 'neighbourhoods' (Can and Heath 2016; Esposito et al. 2020; Hidayati et al. 2021; Ozbil et al. 2021; Zerouati and Bellal 2020) and 'parks' (Chiang and Li 2019; Sheng et al. 2021; Zhai and Baran 2016; Zhai et al. 2018). As such, two distinct categories of syntactic and experiential topics, architectural and urban, are identified and the topic modelling is separately run for both types.

# **Topic modelling**

## Data preprocessing

Data preprocessing uses NLP, a subfield of artificial intelligence (AI), which enables machines to understand and analyse natural language data. There are generally three preprocessing steps in topic modelling, tokenization, stopwords, and stemming (AlSumait et al. 2008; Barde and Bainwad 2017; Syed and Spruit 2017). First, tokenization separates a text into sentences, then into words. That is, it breaks down the given text into the smallest units, so-called "tokens". The present research uses the *NLTK* (Natural Language Toolkit) library package for statistical NLP for English written in *Python* for this purpose.

After tokenizing collected abstracts, the next step is removing stopwords such as "a", "to" and "describe". Whilst the *NLTK* package provides a list of stopwords, the present research uses the stop-words package, which provides common lists of stopwords in various languages in Python, and updates it with one of the most comprehensive collections of stopwords (https://github. com/stopwords-iso/stopwords-en/). Furthermore, general words in Space Syntax research articles (e.g., space, syntax, space, approach, finding, result, etc.) are also excluded. In this way, meaningful words for inferring syntactic and experiential topics remain in the corpus of abstracts. In the stemming process, the tokenized words are further converted and shortened into their root form, removing morphological affixes. For this article, the NLTK Porter stemmer is used for this purpose. For example, stemming "change", "changed", and "changing" returns "chang", "chang" and "chang".

In summary, the collected abstracts are transformed using NLP through these three data preprocessing steps into clean and consistent texts. For example, a sentence, "this paper argues that the evaluation of architectural spaces as numerical entities can identify seemingly random patterns of movement behaviour," is converted to a set of tokens, ['argu', 'evalu', 'architectur', 'numer', 'entiti', 'identifi', 'seemingli', 'random', 'pattern', 'movement', 'behaviour'], which is then fed into the LDA model for Machine Learning (ML).

# LDA topic modelling

LDA, first introduced in 2003 (Blei et al. 2003), is a generative probabilistic model that extracts topics from collections of texts or documents, based on topic probabilities (AlSumait et al. 2008; Blei et al. 2003; Jelodar et al. 2019; Kee et al. 2019) and semantic similarities (Blei et al. 2003). Overcoming the limitations of previous topic modelling methods-such as Latent Semantic Indexing (LSI), Probabilistic Latent Semantic Analysis (PLSA) and Vector Space Model (VSM)-LDA is widely used to capture the heterogeneity of large research topics and to identify a set of core topics using particular words (AlSumait et al. 2008; Jelodar et al. 2019; Syed and Spruit 2017). LDA topic modelling treats a document as a combination of topics (Barde and Bainwad 2017) and each topic as a "multinomial distribution" over the fixed vocabulary or dictionary (Blei et al. 2003; Syed and Spruit 2017). A word (or token in this research) is the basic unit and an item from a vocabulary (Blei et al. 2003) After converting documents to a "bag of words" (BoW) representation using the vocabulary, LDA topic models identify a set of topics consisting of a set of words and their topic proportions, based on probability distribution (Blei et al. 2003). In addition, LDA models were built using the Gensim Python library, which has been frequently adopted for this purpose (Barde and Bainwad 2017; Syed and Spruit 2017). In this way, LDA topic modelling identifies meaningful topical mixtures within the given abstract data. In LDA, the best model (the number of significant topics) could be determined by *perplexity* and *coherence* scores (AlSumait et al. 2008; Blei et al. 2003; Syed and Spruit 2017), which were also used in the present research. In the Gensim library, 'log\_perplexity' is calculated for this purpose. Since it returns per-word likelihood bound, a small 'log perplexity' score implies deterioration. Thus, topic models with bigger scores are selected in the Gensim algorithm. In contrast, a coherence score reveals the degree of semantic similarity between words related to a topic. A higher coherence score generally indicates a better model. There are two coherence measures ( $C_V$  and  $C_{llmass}$ ) frequently used for this purpose (Mimno et al. 2011; Röder et al. 2015; Stevens et al. 2012). The former is based on pointwise mutual information and cosine similarity, while the latter focuses on a word co-occurrence statistic dealing with topic quality (Dahal et al. 2019). The present research uses  $C_{Umass}$  in the Gensim library to validate the quality of topic models (Röder et al. 2015; Stevens et al. 2012).

Another crucial factor is the quantity of iterations (trains). Theoretically, coherence scores increase over the iterations (Bunk and Krestel 2018; Katarya et al. 2022). That is, more is better, but training time increases with the number of iterations. The precise figure varies in applications. For example, Katarya et al. (2022) use 30 iterations, while AlSumait et al. (2008) use 500 iterations. Finally, classifying topics in topic modelling tends to target a small number of topics, often three (Kee et al. 2019; Stracqualursi and Agati 2022; Syed and Spruit 2017). Given the scope of the current research, a small topic number (e.g., k=3 to 6) is both viable and desirable. In

summary, following standard practice in past research (AlSumait et al. 2008; Blei et al. 2003; Syed and Spruit 2017), when determining and validating the best topic model, this research considers perplexity and coherence scores, as well as the number of iterations.

#### Classification and interpretation

The final process of LDA topic modelling is the classification and interpretation of the model outputs. Again, the LDA produces a set of topics with a mixture of words (AlSumait et al. 2008; Jelodar et al. 2019; Syed and Spruit 2017). That is, each topic cluster is labelled by the most frequent words. To do this, the LDA topic modelling produces the three most frequent words per topic. Each word or token is also presented with a respective beta value which is the probability of its occurrence within a topic (Kee et al. 2019). In addition, the final topic model is verified using an inter-topic distance map using principal component analysis and multidimensional scaling using the LDAvis library, because LDA using a Bayesian model is based on random mixtures over latent topics (Blei et al. 2003). Finally, a simple word cloud is presented to help readers intuitively understand topic words in a topic model.

# Results

### LDA topic model training

The architectural corpus for the first LDA topic modelling consists of 27 abstracts. After multiple trials, the LDA topic model training focuses on the optimal topic number (k). First, the number of iterations can be determined with k as illustrated in Fig. 3, where coherence scores ( $C_{Umass}$ ) are used as the index. Unexpectedly, the coherence scores in the figure fluctuated every 5 iterations. Nevertheless, over the first 200 iterations, both k=3 and k=6 displayed relatively higher coherence scores. Using *perplexity* scores for confirmation, the combination of k=3 and 155 iterations was identified as

ideal for the Architectural corpus topic modelling. Using the same process, the urban corpus consisting of 39 abstracts uses the combination of k=3 and 165 iterations for the topic modelling.

# Syntactic and experiential topics *Architectural topics*

Considering LDA's random distribution, the sets of architectural topics generated by LDA were examined with their inter-topic distance maps. After multiple runs were completed, a topic model consisting of three topics was identified (Table 1) and then verified by the inter-topic distance map in Fig. 4. The first topic labelled 'wayfinding behaviour, encompasses the combination of 'wayfinding,' 'behaviour', and 'location', covering 43.1% of tokenized words. Following this same method, the next two topics are 'interactive accessibility' (A2) and 'healthcare design' (A3). Considering the fourth and fifth tokenized words of the second topic such as 'visibility' and 'physical', the second topic encapsulates research about accessibility issues in various spatial units. Since medical space is a popular subject in recent Space Syntax research (Lee et al. 2023), it is not surprising that the LDA topic modelling



Iterations

- k = 5

k = 6

Fig. 3 Coherence scores for the first 200 iterations for the architectural corpus by the number of iterations topic (k = 3 to 6)

	i i		
No.	Topic label	Topic words with a contribution	Token coverage
A1	Wayfinding behaviour	wayfinding (0.019), behaviour (0.012), location (0.012)	43.1%
A2	Interactive accessibility	unit (0.019), interact (0.015), access (0.013)	34.4%
A3	Healthcare design	hospital (0.019), pandemic (0.015), macrocognition (0.013)	22.5%

Table 1 Three architectural topics with topic words



Fig. 4 The inter-topic distance map of three architectural topics

developed a distinct healthcare topic. 'Pandemic' also contributes 15% to the last health-related topic.

As the last step in the LDA process, Fig. 5 is a word cloud developed from the architectural data. Excluding stopwords, the top 10 key words in order of frequency

are wayfinding, design, behaviour, visibility, location, hospital, environment, social, interaction, and patient. These ten words could also encapsulate the most frequent and potentially important keywords in this architectural area of Space Syntax research. They reflect the growing research focus on process, place and people, as key syntactic and experiential factors. For example, wayfinding, visibility, and interaction can be categorised as part of the socio-spatial process dimension. Likewise, location, hospital, and environment belong to the place dimension. Lastly, future socio-spatial research should expand consideration of the people dimension (e.g., patient and visitor). This additional categorisation of socio-spatial factors potentially offers some insights which could be explored in a follow-up study.

### Urban topics

Table 2 identifies three significant urban topics with their top topic words as determined by the LDA topic modelling. Based on topic words, the three topics are 'pedestrian movement', 'park accessibility', and 'cognitive city'. The first observation about these topics is that they are quite different from the architectural ones in Table 1. For example, whilst 'wayfinding' is the most frequent topic word in the architectural data, 'pedestrian movement' is the most significant factor in the urban research. Several



Fig. 5 LDA word cloud developed from the architectural abstract corpus (higher word frequency is reflected in larger words)

No.	Topic label	Topic words with a respective beta value	Token coverage %
U1	Pedestrian movement	pedestrian (0.022), movement (0.014), function (0.010)	35.3
U2	Park accessibility	park (0.032), street (0.018), access (0.011)	33.7
U3	Cognitive city	activity (0.013), city (0.013), cognition (0.012)	31.0

#### Table 2 Three urban topics with topic words

urban topic words also focus on place (e.g., 'park' and 'city') and user (e.g., 'pedestrian'). In this context, the first topic, accounting for 35.3% of topic coverage, represents the combination of 'pedestrian', 'movement' and 'function'. The second topic highlights 'park accessibility', including 'park trail' or 'urban trail'. In contrast, the last topic focuses on socio-spatial activities to improve the image of a city.

Figure 6 presents a word cloud derived from the abstracts of the urban articles, identifying the 50 most frequent specific words. The top 10 words in order of frequency are street, park, pedestrian, social, public, city, accessibility, behaviour, walkability, and movement. The urban research was focused on both environmental and socio-spatial factors, which were also captured by the LDA (Table 2), in the 'place' dimension (street, park, and city). This result is noteworthy because the architectural research highlights 'wayfinding' and 'interaction'. For example, 'accessibility,' walkability', and 'movement' are addressed in the data. Collectively, Fig. 6 indicates the complexity of spatial experiences in urban environments, which is discussed in the following section.

#### Space Syntax topics and their implications

Since the syntactic and experiential topics identified in this article have multiple implications for current and future research, each topic generated by the LDA topic modelling is further discussed with relevant articles in the collected data and key references.

#### A1. Wayfinding behaviour

Research on wayfinding behaviours is dominant in the architectural data. First, holistically investigating vertical and multilevel wayfinding behaviours, Hölscher et al. (2012) confirm the predictive power of syntactic measures such as connectivity and integration. Although their study reviews past wayfinding experiments and performance measures, the combined route-based approach of ALA and VGA was used to distinguish between experienced and inexperienced participants' wayfinding performances. Li and Klippel (2012) also investigate the relationship between spatial configurations and wayfinding behaviours in a multiple-level building and echo the view that syntactic measures such as visibility (VGA) and connectivity (ALA) are effective for examining human wayfinding performance. In contrast, Nubani et al. (2018) carefully validate the relationship between VGA-focused data (visual integration and isovist area) and the number of visits a person makes to a space. Specifically, they measure the visual saliencies of specific locations using an object-based salience rating. In this way, affirming general socio-spatial phenomena in past studies, they validate individual cognitive mapping assumptions in Space Syntax research. De Cock et al. (2022) also investigate individual spatial cognitions by VGA and isovist measures. Interestingly, unlike conventional Space Syntax research on wayfinding, they focus on the differences in wayfinding and gaze behaviours in a virtual reality (VR) environment to reveal the impact of route instruction types on cognitive load or attention. Recent wayfinding research in architecture, like these examples, is often focused on individual spatial cognition processes. Along with individual experience, visual salience, signage, and sense of direction should all play a role in influencing wayfinding performance.

To compare wayfinding behaviours in Chinese and Australian hospitals, Geng et al. (2021) adopt a visibility-focused approach using a combined VGA, ISA, and ABS method Their observations include consideration of building typologies and design elements such as



Fig. 6 LDA word cloud developed from the urban abstract corpus (higher word frequency is reflected in larger words)

colour-coded floors, family-friendly design, and natureintegrated environments. Their study is, however, limited to differentiations of syntactic variables using descriptive observations. This type of comparative research is useful in the Space Syntax community, but it would be more valuable if conducted in conjunction with the appropriate experiential variables identified in this review. In contrast, Tzeng and Huang (2009) conduct an in-depth investigation into behavioural (e.g., stop, search, decide, and legibility) and visual (e.g., signage) content for wayfinding design. Specifically, in their research data are analysed in terms of wayfinding frequency and distribution. The categorisation of wayfinding content is also applicable to future behavioural research in the built environment. However, rather than the relationship between socio-spatial variables, their analyses are focused on the individual influence of spatial, behavioural and visual factors on wayfinding. Interestingly, these two wayfinding articles dealing with 'medical space' include a greater level of consideration of design elements than the wayfinding research in the rest of the 'architecture' set of articles. Recently, Khotbehsara et al. (2023) investigate the impact of the COVID-19 pandemic on visitors' wayfinding behaviours in hospitals, using a hybrid method of syntactic measures and simulations (ALA, ABS, VGA, and ISA) as well as diverse empirical data (gate counting, people following, and interview). A comparison between the collected data before and after the pandemic strongly suggests the need for accessibility and legibility in postpandemic healthcare centres.

#### A2. Interactive accessibility

In terms of 'social interaction', the research of Ferdous and Moore (2015) presents a rigorous 'spatial behaviour interaction model' consisting of spatial variables (visibility and proximity) and behaviour (visible co-presence, movement, and interaction), which might be major socio-spatial properties in Space Syntax research. Specifically, their two levels (low and high) of social interactions are useful for interpreting the interactional outcomes of architectural configuration. However, the use of VGA measures in isolation might be limited to comprehensively investigating people's behaviours with dementia. Büyükşahin (2023) also uses VGA to examine the impact of the COVID-19 pandemic on spatial preferences and behaviours in shopping malls, focusing on humanspace interaction. The results of survey discover some changes in spatial preferences and usage habits after the pandemic, but the relationship between spatial configuration and human behaviour is rarely discussed in the article. In contrast, O'Hara et al. (2018) present an ISAfocused approach to examine situated macro-cognitive interactions in a healthcare environment, providing an in-depth analysis of observation and focus group data. Furthermore, their study explains formal and informal interactions with Space Syntax constructs (not syntactic measures) developed from the human data analysis, which would be useful for examining more qualitative characteristics in the built environment. This focused ethnographic approach could offer an alternative to experiential data analysis in terms of qualitative evidence, that is rarely applied in Space Syntax research. In contrast, Zeng et al. (2020) develop a novel framework to investigate physical and psychological changes in culturally initiated spatial regeneration. Specifically, a bilinear measurement model of acculturation is used to quantify complex cultural identity, while a JPG is used to measure spatial transformation and develop architectural genotypes. In this way, the relationship between spatial and cultural data is presented in their study.

'Interactive accessibility' is also associated with environmental psychology. For example, Keszei et al. (2019) examine the capacity of Appleton's prospect-refuge theory to capture the causes of different interactive behaviours in the built environment. While Space Syntax methods have been used in environmental psychology, the comparison of two distinct types of social situations (prospect-demanding and refuge-demanding) is novel. However, their simulated case study using VGA is based on a simple spatial program and the impacts of designensembles in the 3D space on syntactic measure and seating choice are not clearly discussed. Visibility should be a major socio-spatial factor in architectural research, but environmental factors (e.g., nature and ceiling height) might be more important psychologically, than simple spatial layout. For example, Yaseen and Mustafa (2023) present a syntactic study on biophilic design parameters in school buildings, highlighting the visibility of nature-connectedness. Although their research is based on VGA, it develops new parameters such as visibility of nature, visual permeability, and naturalness of view. In this way, spatial syntax can be complementary to 'nature syntax' in design evaluation.

Particularly, 'accessibility', a fundamental Space Syntax issue in both architectural and urban studies, is often closely related to the other topics. For example, it can be linked to 'A1. wayfinding behaviour' as well as healthcare research. Omer and Goldblatt (2017) use ALA and SA to investigate the structure of movement patterns and flows in two shopping malls. They adopt Q-analysis as a segment-based approach, capturing connectivity as well as topological and metric mean depths. Consequently, their combined approach has the potential to reveal both individual movement paths and their aggregate patterns at the same time, which are rarely addressed in past research. Q-analysis is a complementary technique

to traditional SA and eccentricity measure is a promising development. Their applications and interpretations in socio-spatial research do, however, require further clarification and assessment. In addition to the use of axial line-based approaches, ABS can be a useful technique for this purpose. For example, Kim and Kim (2023) reveal that ABS can explain pedestrian interactive behaviours seeking the shortest paths between ticket gates and station entrances in a subway station. In contrast, Aknar and Atun (2017) present a unique approach to revealing 'sub-integration' locations that are neither integrated nor segregated spatially. Their approach using VGA is, however, limited to one dimension of syntactic variables, visual integration, disregarding other socio-spatial factors like spatial function. In addition, although their use of Fibonacci retracements can improve the predictive capacity of syntactic results, spatial selection behaviours (choosing seating, for example) might be more impacted by other design elements such as light, material, and furniture. In contrast, ALA using a fully furnished plan layout could provide a complementary approach to this problem. These new topological measures over traditional approaches might be valuable, but the definitions and interpretations of various measures might be ongoing issues in the Space Syntax communities.

### A3. Healthcare design

In addition to wayfinding behaviours (Geng et al. 2021; Tzeng and Huang 2009) in hospitals, discussed in the first topic (A1), the architectural studies dealing with medical spaces have multiple ramifications for socio-spatial interaction or usability in healthcare facilities (Ferdous and Moore 2015; Li et al. 2022; Neo and Sagha-Zadeh 2017; O'Hara et al. 2018; Rashid et al. 2014, 2016). Rashid et al. (2014), for example, explore behavioural and psychological responses in ICUs using a combined method of ALA and JPG, presenting statistically significant correlations between interactive behaviours and syntactic properties. Although there is a limitation arising from the data collection method, Rashid et al's work offers a cogent description of the socio-spatial factors in the built environment. The discussion, however, about different types of integrations (axial and node) and their relationship with different behaviours could be further articulated using Hillier's (1999) line and topological representation of space. Researchers also suggest that the use of snapshot observations (Askarizad and Safari 2020; Can and Heath 2016; Zerouati and Bellal 2020) could capture additional interaction behaviours. In a similar study, Rashid at al. (2016) highlight physical (ALA) and visual (VGA) accessibilities to investigate staff perception and interaction in ICUs. Interestingly, they find that axial integration might be less important than visual integration in predicting social interactions. This result also reinforces the decision, seen in many architectural studies in this review, to use VGA to examine socio-spatial factors. However, although experiential data is collected from both a survey and observations, Rashid at al's (2016) regression analysis is limited to the effect of syntactic properties on interaction frequency and length. The quality of interaction or behaviour is not addressed. Ferdous' and Moore's study (2015) and O'Hara et al.'s research (2018) discussed in the previous topic (A2. Interactive accessibility) are both pertinent for this purpose. In contrast, Neo and Sagha-Zadeh (2017) measure visibility and global traffic flow scores based on Space Syntax theory, to optimise the locations of hand sanitising stations (HSS) in healthcare environments. Specifically, they present statistically significant regression models using electronic tracking data. Recently, in response to the COVID-19 pandemic, Li et al. (2022) examine the social organisations of Fangcang shelters (large-scale, temporary hospitals in China), using interview and social media data. Even in the most isolated spaces-akin to 'total institutions'-spatial structures have an impact on patients' collective interactions and emotional stimulation.

Architectural research on healthcare design also addresses specific users' behaviours and perceptions (Alalouch and Aspinall 2007; Cai and Spreckelmeyer 2022; El-Hadedy and El-Husseiny 2022; Joshi et al. 2023; Nicoletta et al. 2022). The research of El-Hadedy and El-Husseiny (2022), for example, examines the spatial configuration of a healthcare facility, with a focus on sites where workplace violence has occurred. They identify crime prevention through environmental design (CPTED) features as important socio-spatial factors that influence human behaviours. Cai and Spreckelmeyer (2022) analyse nursing unit designs using a combination of VGA and ISA, comparing with the post-occupancy evaluation (POE) results of hospital patients. However, in both articles, experiential data-field observations, interviews, and POE-are only treated descriptively and subjectively with reference to syntactic measures. In contrast, comparing six wards, Alalouch and Aspinall (2007) use VGA to statistically reveal that the general perception of private locations in plan layouts is positively related to visual integration. Despite the high levels of significant differences in their study, its generalisability remains unclear. Furthermore, the study did not consider users in hospitals or 3D environmental factors, both of which would affect the results. Nicoletta et al. (2022) investigate the spatial layouts of two maternity unit settings using VGA, ISA, and JPG, identifying relationships between syntactic data and mothers' and midwives' perceptions. Likewise, Joshi et al. (2023) examine the relationship between visual exposure at handoff locations (using isovist connectivity) and physicians' perceptions (interruptions, privacy, and collaboration). Acknowledging contextual variables and other factors impacting on their results, they demonstrate significant differences in the socio-spatial characteristics of three distinct physician workstations.

#### U1. Pedestrian movement

Pedestrian movement in urban space is, obviously, a dominant socio-spatial focus of the collected urban data. For example, Yang and Vaughan (2022) investigate how pedestrian movement is impacted by street configuration (centrality), functional uses, and physical structures (e.g., land number, sidewalk, and tree), before they reveal significant differences between the effects of gated and non-gated housing estates. Yıldırım and Çelik (2023) also explore important determinants of pedestrian movement in historical centres through VGA, behavioural mapping, and an evaluation survey (sense of place and content). Specifically, density, activity, and mobility maps are generated by pedestrian counting and observation. Based on Gordon Cullen's theory of pedestrian perception, it is clear, both quantitatively and qualitatively, that spatial configuration influences pedestrian movement and behaviour. Likewise, Domènech et al. (2020) reveal the relationship between spatial characteristics and tourists' behaviours, identifying several indicators for specific behaviours. In addition to ALA measures, they calculate other factors-mobility, physical attributes, commercial activity, and visibility—which affect walkability in a city. GPS data is used for measuring mobility, and Terrain Elevation Model, Laser Imaging Detection and Ranging (LiDAR) and a field survey are used to develop physical attributes (slope, vegetation, and maximum car speed, respectively).

Visibility properties of points of interest (POIs) are manually measured in the field, but VGA or ISA might also be useful for this purpose. Collectively, Domènech et al's multi-dimensional framework is clearly verified with an extensive data set and regression analyses, whereas their time-based mobility and topographical elements require further clarification. In contrast, Koohsari et al. (2016) address associations between walkability and syntactic indexes. A walkability index is calculated using residential density, intersection density, retail floor area ratio and land use mix, while a Space Syntax walkability index is developed from gross population density and the integration of ALA. In this way, Koohsari et al. reveal the impacts of both walkability indexes on walking, although their findings rely on self-reported measures. While these studies focus on walkability, their ALAs notably include motorways as part of their maps. Such methodological limitations often exist in Space Syntax research and while they are common, they should be addressed by future researchers.

Urban studies frequently examine the impact of syntactic values (shortest path and least directional change) on pedestrian route choice (Chiang and Li 2019; Hillier 1999; Hillier and Iida 2005; Jiang 2009; Karthika et al. 2022; Zhang and Chiaradia 2022), which is closely related to wayfinding research. The research of Ozbil et al. (2021), for example, reveals a strong correlation between street network connectivity (metric reach and directional reach) and children's active school travel. They further measure other factors such as land use, street design and streetscape, some of which are related to walkability to and from school. Like Ozbil et al's study (2021), Shatu et al. (2019) also collect street characteristics (sidewalk, land use, and traffic) and pedestrian travel routes using a self-reported questionnaire. Their findings suggest that two syntactic values, distance and direction, could explain more than 50% of route choice, which is in line with past Space Syntax results. At the same time, they indicate the importance of other environment factors in creating a pedestrian-friendly urban design. In contrast, Mansouri and Ujang (2017) also reveal that the diversity of land uses and street activities are more influential than street connectivity. Furthermore, they find no, or negligible, correlations between pedestrian intensities and integration values of multiple POIs. Although their study is limited to historical districts, gate observation and integration measured from ALA, the results imply that multi-dimensional factors collectively impact on walking behaviours in an urban system. From this point of view, Hidayati et al. (2020) discuss safe and inclusive environments through a combined analysis using ALA and SA. They also highlight that pedestrian accessibility is related to land uses and inhabitants' perceptions of walking environments. Except for a correlational analysis of school locations and vehicular or pedestrian throughmovements, however, their study is largely descriptive. Nonetheless, it provides insights into how Space Syntax research can address the dynamic socio-spatial factors of transport, land use and social integration in urban planning.

Recent research interests in pedestrian movement include those examining spatial vitality in urban underground space (Xu and Chen 2022), quantifying visuofunctional spaces (Shen and Wu 2022), and comparing 3D multi-level pedestrian networks (Zhang and Chiaradia 2022). These studies provide unique analytic or syntactic approaches. For example, Xu and Chen (2022) highlight the spatio-temporal use of urban underground spaces, along with two syntactic measures such as accessibility and visibility. They also consider the relationship between spatial vitality (cross-sectional pedestrian flow) and physiological environment (illumination, temperature, and wind speed). Shen and Wu (2022) present a new version of VGA, functional visibility graph analysis (FVGA), which develops three basic measures (visible function size, entropy, and mean angular depth step) and advanced ones (visual function connectivity, function visibility, and visible function closeness). Using social media check-in data, the FVGA model at the pedestrian scale precisely reveals the short-term transformation of the visual landscape. Lastly, Zhang and Chiaradia (2022) not only extract 3D indoor and outdoor multi-level pedestrian networks, but also use a 3D hybrid angular-Euclidean analysis. In this way, they overcome some conventional criticisms of 2D axial line models and improve pedestrian route choice, wayfinding, and movement research.

#### U2. Park accessibility

Accessibility of, or movement through, parks has also been widely explored in past research. First, in terms of accessibility, the research addresses topological proximity to parks. Chiang and Li (2019) examine the metric or topological proximity, its relationship to the frequency of visits to parks, and to perceived stress. Topological proximity is a significant predictor of visits and stress, while their study finds metric proximity less effective for this purpose. This result confirms past research (Hillier 1996; Shatu et al. 2019), while some other factors identified as similarly significant-qualities of parks, attractions, and environmental designs-are not considered in Chiang's and Li's study. For example, urban density and visibility would also impact on perceived urban stress (Knöll et al. 2018). Safaie Ghamsary et al. (2023) also measure the impact of land use and accessibility on public presence in a historical neighbourhood to effectively locate pocket parks. ALA (connectivity, integration, and mean depth) is used to investigate residents' access in the neighbourhood, while a survey develops land use and accessibility indices. Likewise, Can Traunmüller et al. (2023) examine park accessibility through the combined approach of ALA and SA. Particularly, their study presents multiple regression analyses to examine how the number of park users is impacted by syntactic data (e.g., integration and choice) as well as various ethnographic data (e.g., population size, facility score, and park safety) collected from field observations.

In contrast, Zhang et al. (2020) investigate visitors' movements within a park. Their spatio-temporal measurement using GPS tracking data is similar to Domènech et al.'s study (2020), but Zhang et al. use a VGA-only approach with a convex map of the park, instead of ALA. In theory, this approach to walkability could connect to wayfinding research in architectural space. Other socio-spatial factors such as individual experience, signage, and sense of direction (Nubani et al. 2018; Tzeng and Huang 2009) could also be examined in future research. In addition, a combined ALA and VGA approach for this purpose could be beneficial. Zhai and Baran (2016) also use a modified CSA but employ graphbased syntactic measures. Interestingly, they segment pathways in a park which are then regarded as nodes in a justified graph (syntactic measure) as well as observation points (experiential measure). This pathway segment approach could be an alternative to the conventional street segmentation approach (Jiang and Claramunt 2004; Jiang and Liu 2009). In addition to syntactic variables, three environmental factors-pathway connection with activity zones, presence of shade, and lateral visibility-are identified as potential predictor variables in their regression model. Although their findings seem to extend beyond the statistical results, their study contributes to an understanding of the relationship between seniors' walking behaviours, syntactic properties, and microenvironmental characteristics. Like Zhai's and Baran's research (2016), their results may be acceptable in Space Syntax research because human behavioural research can present  $R^2$  values lower than 0.5 and still be useful. In this context, these walkability studies in parks reiterate the impact of both spatial configuration and attraction on natural human movement in an urban system (Hillier et al. 1993) closely related to the first urban topic (U1).

Lastly, urban accessibility research, typically based on ALA, is closely related to this topic, but it deals with specific sites or destinations, e.g., religious sites (Karthika et al. 2022), heritage buildings (Hegazi et al. 2022), playgrounds (Istiani et al. 2023), public libraries (Zhao and Hong 2023), and resettlement communities (Yang et al. 2023). In particular, Karthika et al. (2022) develop a spatial time-based accessibility map using ArcGIS network analysis and video data for crowd management, while Hegazi et al. (2022) use a combination of syntactic measures (ALA and VGA) and behavioural maps to investigate socio-spatial vulnerability assessment. Addressing the pandemic and post-pandemic situations, Istiani et al. (2023) also generate occupancy and observation maps for this purpose. Yang et al. (2023) distinguish physical (or syntactic) accessibility from spatial accessibility that is regarded as a household's ability to access various services, but integrate both to establish indices of multiple deprivations for sustainable urban development. Conversely, Liu et al. (2018) focus on the scenic kernel density values of mountain areas, revealing the relationship between syntactic values (axial integration and control value) and tourist trail flows. Similarly, Zhai et al. (2018) combine a CSA derived, stroke-based network of trail segments in forest parks, with human data collected from GPS trackers. They also confirm a socio-spatial relationship between syntactic properties and environmental characteristics, but their regression model does not fully account for several aspects of the human data. Tourism attraction is conceptually linked to other urban topics such as 'pedestrian movement' and the 'cognitive city', dealing with scenic and environmental factors.

#### U3. Cognitive city

Aside from research about pedestrian and spatial networks, a growing research area in the Space Syntax community is concerned with the perceptual characteristics of a city. As an example, Elshater et al. (2019) use ALA data to compare two Egyptian cities' singularity characteristics, presenting the interconnectedness of physical and non-physical urban design concepts. In that research ALA is applied to capture people' perceptions through movement. Their study, however, doesn't discuss the relationship between the urban singularity elements (e.g., architectural style, urban artefacts, human experience, etc.) and syntactic measures (connectivity, choice, integration, intelligibility, and synergy), but separately presents their values, before revealing the differences between the two cases. Pan et al. (2022) also use ALA to examine perceptions of location-specific urban spaces, focusing on accessibility, choice, and public activity, to correlate urban morphological data (accessibility and choice) and human distribution data. Interestingly, Wang et al. (2022) use a human-machine adversarial model over the use of Space Syntax, based on deep learning, to predict residents' perceptions of city streets. They also present automatic image semantic segmentation and visual element classification, which are recently used for computer vision or machine learning. As such, they not only measure street accessibility, but also evaluate the quality of streetscapes or street view imagery (SVI). In this way, their study enables the precise investigation of urban spatial perception both horizontally and vertically.

Güngör and Harman Aslan (2020) revisit Lynch's theory of urban visual perception with a syntactic approach combining ALA, SA, and VGA. Their research provides a precise syntactical interpretation of Lynchian urban elements. Experiential data is, however, developed from observation, questionnaire, interviews, and cognitive mapping and are only used for descriptive analysis. Nonetheless, the conceptual mixture of cognitive and syntactic urban design elements is promising. Furthermore, cognitive mapping in their study should not only be applicable to wayfinding research, but also useful for improving a city's imageability. Bai et al. (2023) further develop a semi-structured interview based on the Lynchian paradigm. They empirically investigate public spaces in historic villages using three aspects of socio-spatial factors: (i) morphological (spatial network centralities using bipartite graphs), (ii) cognitive (cultural significance and space usage) and (iii) behavioural (Wi-Fi positioning tracking). The theory of imageability has often been discussed in Space Syntax research (Askarizad and Safari 2020; Elshater et al. 2019; Li and Klippel 2012), and especially in terms of spatial cognition (Kim and Penn 2004; Marcus et al. 2016). Esposito et al. (2020), for example, use spatial cognition to explore wayfinding route choices associated with landmarks. Using a combination of quantitative ABS and qualitative data, their research confirms that syntactic values can be predictive of wayfinding and environmental perception. The impact of environmental factors, e.g., visual element proportion (Wang et al. 2022) and the quality of landscape elements (El-Darwish 2022) are, however, not fully explained in the syntactic analysis. In addition, socio-economic determinants like safetysee El-Hadedy's and El-Husseiny's study (2022) or Can Traunmüller et al.'s work (2023)-can be considered in spatio-cognitive research of this type. Recently, cognitive mapping combined with ALA is used to investigate suburban village tourism (Zheng et al. 2022) and VR environment navigation (Brunec et al. 2023).

With parallels to 'A2. Interactive accessibility' in the architectural studies, 'urban interaction' is an important socio-spatial issue in the cognition and design of urban morphology. Unlike the architectural studies, which largely relied on VGA, urban studies often used ALA to examine interactional frequencies or locations, regardless of the qualities or types of social interaction. Only Sheng et al. (2021) consider two types of urban interactions, personal (or intimate) and social groups in urban parks, measuring the intensity of social interaction by the number of observed people. Furthermore, their syntactic measure is based on a pathway segment approach, similar to Zhai's and Baran's (2016), but focuses on angular integration and choice via SA. Along with syntactic properties, they confirm that spatial factors such as pathway length and zone area are positively associated with social behaviours. In contrast, Can and Heath (2016) address the frequency of urban interaction using a questionnaire and the location and number of groups recorded in snapshot observations, but never examine the type of interaction. Results confirm that the syntactic measures (integration, connectivity, synergy, and intelligibility) developed from ALA are acceptable predictors for movement, co-presence, and social interaction. Zerouati and Bellal (2020) also use a snapshot observation approach, but develop snapshot maps with VGA results that include the types of activities and users. In this way, they provide sound visual evidence about socio-spatial relationships. Similarly, Gümüş and Yılmaz (2022) develop behavioural maps, which are compared with syntactic data such as

integration and connectivity. Askarizad and Safari (2020), instead, employ four observational approaches, including snapshot observations to capture static activities, which are only used for descriptive relationships alongside syntactic results. Furthermore, urban interactions in their study are not measured, instead they are qualitatively assessed using behavioural patterns in public spaces. ALA, ISA and VGA provide important design knowledge about urban morphology and social relations, whereas attractions and environmental factors could be further explored. Walkability research (Zhai and Baran 2016; Zhai et al. 2018) provides a useful reference in this regard.

A small number of studies address safety and social integration in urban spaces, which might be related to the development of a city's imageability. Compared to Chiang's and Li's proximity research (2019), Knöll et al. (2018) consider both syntactic (integration and visibility) and environmental properties (density and spatial typology) in their regression model of perceived urban stress. Specifically, they reveal that open space typologies, validated by a survey of environmental characteristics, are the best predictors and are superior to syntactic values in this regard. This result articulates the importance of environmental factors in Spacy Syntax research. In contrast, Mohamed and Stanek (2020) focus on the conventional socio-spatial relationship between syntactic and experiential data, disregarding the role of other determinants such as economic and environmental characteristics. In terms of the frequency of sexual harassment incidents, Mohamed and Stanek (2020) find the quality and type of physical surroundings may be influential factors in their regression model. While there are precedents for this type of study of the impact of spatial integration on crime (El-Hadedy and El-Husseiny 2022), they have not received the same level of investigation in the Space Syntax community. Furthermore, there might be socio-cultural factors shaping this result.

#### Summary

This section summarises the ramifications of the research and the collected data. The 27 architectural studies explore planning in architectural spaces, although there is also a clear emphasis on medical or healthcare spaces with 16 articles having this focus. The eleven nonhealthcare studies identify important considerations for improving understanding of wayfinding behaviours in complex buildings (De Cock et al. 2022; Hölscher et al. 2012; Li and Klippel 2012; Nubani et al. 2018) and interactive behaviours (Aknar and Atun 2017; Büyükşahin 2023; Keszei et al. 2019; Kim and Kim 2023; Omer and Goldblatt 2017; Yaseen and Mustafa 2023; Zeng et al. 2020). Space Syntax approaches to investigating healthcare facilities have become relatively common in the last decade (Haq and Luo 2012; Lee and Ostwald 2020; Lee et al. 2023), often examining syntactic and experiential issues using hybrid approaches.

In contrast, the 39 urban studies collected in this review, which generally focus on socio-spatial interactions in streets and parks, potentially have significant implications for urban planning and policymaking (Chiang and Li 2019; Lee et al. 2023). By examining the relationships between syntactic properties and experiential values, Space Syntax research can provide quantitative knowledge about pedestrian movements (Domènech et al. 2020; Koohsari et al. 2016; Mansouri and Ujang 2017; Ozbil et al. 2021; Shatu et al. 2019), design for accessibility and use of parks (Can Traunmüller et al. 2023; Chiang and Li 2019; Liu et al. 2018; Safaie Ghamsary et al. 2023; Zhai and Baran 2016; Zhai et al. 2018; Zhang et al. 2020), and policies and strategies to improve the image of a city (Bai et al. 2023; Elshater et al. 2019; Esposito et al. 2020; Güngör and Harman Aslan 2020; Wang et al. 2022), which also shapes social interactions (Askarizad and Safari 2020; Can and Heath 2016; El-Darwish 2022; Sheng et al. 2021; Zerouati and Bellal 2020) as well as safety and social integration (Hidayati et al. 2021; Knöll et al. 2018; Mohamed and Stanek 2020).

#### Discussion

This review has used a new combined approach of PRISMA and LDA to scope and partially synthesise a large body of research, offering the opportunity to undertake a detailed critical review of the content of this work. The combination of PRISMA and LDA is, in the field of architecture, a clear methodological innovation. The results of LDA topic modelling were not only able to capture a more nuanced and deeper set of keywords and concepts, but also their semantic relationships. Significantly, LDA also provided a measure of the precise distinction between topics, something other methods cannot do. In addition to the perplexity and coherent scores, the intertopic distance map using multidimensional scaling (see Fig. 4) was used to validate the topic models. Two word clouds (Figs. 5 and 6) also confirmed the automaticallygenerated topics. However, despite these validations and verifications, the final set of topics might be limited in its reliability. In particular, the random topic generations could be problematic. For example, each run of LDA was slightly unpredictable, producing similar (but not identical) sets of topics. Labelling and interpreting the generated topics was also dependent on the researchers' knowledge and experience. This is one of the reasons why the extensive analysis of the topic literature was undertaken for this article, to 'unpack' the connections made by LDA. An alternative approach, that would need to be

carefully tested, would be to use full-text data rather than abstract data for LDA (Syed and Spruit 2017). Future research using LDA should address these challenges and limitations.

(2024) 11:3

#### Research gaps and future research directions

Lee and Ostwald City, Territory and Architecture

As shown in the LDA topic models, it is apparent in the present research that human experience or behaviour in the built environment is a multi-dimensional issue. That's why socio-spatial, environmental theories-for example, Lynch's theory (Güngör and Harman Aslan 2020) and Appleton's theory (Keszei et al. 2019)-have frequently been examined in Space Syntax research. Additional psychological and cognitive theories could also be included to understand the complexity of socio-spatial experience. However, the psychological impact of the built environment, like 'perceived stress' (Knöll et al. 2018), is largely an uncharted area. Thus, such theories should be further examined with existing syntactic measures for a range of architectural and urban spaces. Likewise, since it might be hard to control socio-economic factors (Sohn 2016), the impact of spatial permeability on crime remains unclear. That is, the combination of Space Syntax and CPTED (El-Hadedy and El-Husseiny 2022) requires further investigation.

Since the early archaeological applications (Hillier and Hanson 1984), archaeological research using JPG has grown, e.g., Pataraya and Wari provincial administration in Nasca (Edwards 2013) and Homol'ovi I, Hopi archaeological site in Arizona (Fladd 2017). However, the conventional applications are limited to syntactical approaches, visualising and estimating ancient socio-spatial relationships. In contrast, diverse spatial experiences in historical sites have also been significantly explored in terms of tourists' movement (Mansouri and Ujang 2017), pedestrian movement and behaviour (Yıldırım and Çelik 2023), residents' access and land use (Ghamsary et al. 2023), and public space with cultural significance (Bai et al. 2023). As such, Space Syntax research on heritage buildings and sites can be expanded to examine their impacts on current spatio-cultural experience as well as to provide spatial resources for future heritage preservation or urban development. On the other hand, spatiocultural experience should be one of the most important research directions in this regard. For example, along with syntactic measures, Zeng et al. (2020) suggest spatial regeneration and acculturation considering the relationship between spatial and cultural properties. Although visitors' and residents' diverse socio-spatial experiences are not considered in their research, a computational approach to cultural and experiential dynamics could be beneficial for future research. Such experiential studies are useful for future Space Syntax research considering socio-spatial behaviours in the built environment. But in many of these cases, there is a clear need to balance qualitative and quantitative approaches due to the complexity of socio-spatial factors.

As Hillier (1996) highlights diverse socio-spatial topics from archaeological sites to hospital designs, various research topics related to healthcare designs are discussed in the last architectural topic (A3). Indeed, research about medical space is categorised into one independent, dominant topic in the collection of architectural literature (Ferdous and Moore 2015; Li et al. 2022; Neo and Sagha-Zadeh 2017; O'Hara et al. 2018; Rashid et al. 2014, 2016). Specifically, recent research addresses the impact of the COVID-19 pandemic on socio-spatial experience. For example, the pandemic has influenced people's spatial preferences and usage habits in shopping malls (Büyükşahin 2023) and visitors' wayfinding in healthcare centres (Khotbehsara et al. 2023). Responding to the pandemic, the isolated space of Fangcang shelters (Li et al. 2022) and the spatial network of playgrounds (Istiani et al. 2023) are also investigated in terms of spatial, social and emotional ramifications. Certainly, the COVID-19 safety measures and physical distancing have affected social interaction and spatial preference, challenging the traditional socio-spatial relationships in the built environment. Although the fundamental socio-spatial principles of Space Syntax theory are still valid, due to its simplicity and abstraction, this kind of social phenomenon, along with economic and environmental factors, should be further addressed in future research.

Lastly, as discussed in the final urban topic (U3), research in spatial cognition has identified some important socio-spatial factors that influence human experiences in the built environment (Dalton 2001; Freksa 2004; Freksa et al. 2007). The complexity of spatial cognition is also not a new topic in Space Syntax (Bafna 2003; Hillier 2012; Penn 2003). Nonetheless, the spatio-temporal nature of a space (Freksa et al. 2007; Netto 2016), and its recursive or interactive relationships (Freksa 2004; Netto 2016; Raban 1974), are still not well addressed in Space Syntax communities and should be important topics for future research.

#### Conclusion

Space Syntax has evolved over almost 40 years to overcome its limitations in architectural and urban analytics and to be accepted as a standard method in many fields. Notwithstanding this status, the understanding of current and new syntactic and experiential topics should remain a priority. This research provides the first probabilistic models for Space Syntax studies on spatial experiences using LDA. Acknowledging the limitations of LDA topic modelling and sampling, this review rigorously develops eleven deep syntactic and experiential topics and qualitatively examines their implications in architectural and urban research, producing new integrated knowledge. The six topics identified in this research—A1. Wayfinding behaviour, A2. Interactive accessibility, A3. Healthcare design, U1. Pedestrian movement, U2. Park accessibility, and U3. Cognitive city—not only have implications for future research, but they also reflect important or timely issues. The application of PRISMA was limited to data collection in this research, but it could be extended to the synthesis of comparative variables.

The combination of PRISMA and LDA is a new and effective way to systematically map the collected data to a set of topics. In addition to the perplexity and coherence scores considered in the topic models, the findings are validated and generalised with the collected literature. This article contributes to identifying key concepts as well as examining the complexity of socio-spatial issues in Space Syntax research, in addition to having methodological significance in a broader range of research areas. A future study will be a full systematic review synthesising heterogenous syntactic and experiential variables through PRISMA.

#### Supplementary Information

The online version contains supplementary material available at https://doi. org/10.1186/s40410-023-00223-3.

Additional file 1: Table S1. 66 Articles included in the data collection.

#### Author contributions

Conceptualisation, methodology and analysis, JHL and MJO; LDA modelling and visualisation, JHL; writing—original draft preparation, JHL; writing—review and editing, JHL and MJO; funding acquisition, JHL and MJO. All authors read and approved the final manuscript.

#### Funding

This research was supported by the ARC (DP220101598) and UNSW Scientia program.

#### Availability of data and materials

The data presented in this research are available from the corresponding authors upon reasonable request.

#### Declarations

#### **Competing interests**

The authors declare that they do not have any competing interests.

Received: 15 March 2023 Accepted: 27 December 2023 Published online: 09 January 2024

#### References

- Abshirini E, Koch D (2016) Rivers as integration devices in cities. City Territ Archit 3(1):1. https://doi.org/10.1186/s40410-016-0030-4
- Aknar M, Atun RA (2017) Predicting movement in architectural space. Archit Sci Rev 60(1):78–95. https://doi.org/10.1080/00038628.2016.1254594
- Alalouch C, Aspinall P (2007) Spatial attributes of hospital multi-bed wards and preferences for privacy. Facilities 25(9/10):345–362. https://doi.org/10. 1108/02632770710772450
- AlSumait L, Barbará D, Domeniconi C (2008) On-line Ida: adaptive topic models for mining text streams with applications to topic detection and tracking. In: 2008 eighth IEEE international conference on data mining, IEEE, pp 3–12
- Askarizad R, Safari H (2020) The influence of social interactions on the behavioral patterns of the people in urban spaces (case study: the pedestrian zone of Rasht Municipality Square, Iran). Cities 101:102687. https://doi. org/10.1016/j.cities.2020.102687
- Asmussen CB, Møller C (2019) Smart literature review: a practical topic modelling approach to exploratory literature review. J Big Data 6(1):93. https://doi.org/10.1186/s40537-019-0255-7
- Bafna S (2003) Space syntax: a brief introduction to its logic and analytical techniques. Environ Behav 35(1):17–29. https://doi.org/10.1177/00139 16502238863
- Bai N, Nourian P, Pereira Roders A, Bunschoten R, Huang W, Wang L (2023) Investigating rural public spaces with cultural significance using morphological, cognitive and behavioural data. Environ Plann B Urban Anal City Sci 50(1):94–116. https://doi.org/10.1177/23998083211064290
- Barber J (2023) Latent Dirichlet Allocation (LDA) with Python. https://rstudiopubs-static.s3.amazonaws.com/79360\_850b2a69980c4488b1db95987 a24867a.html. Accessed 23 Feburary 2023
- Barde BV, Bainwad AM (2017) An overview of topic modeling methods and tools. In: International Conference on Intelligent Computing and Control Systems (ICICCS), 15–16 June 2017. pp 745–750. https://doi.org/10. 1109/ICCONS.2017.8250563
- Benedikt ML (1979) To take hold of space: isovists and Isovist fields. Environ Plan 6(1):47–65. https://doi.org/10.1068/b060047
- Blei DM, Ng AY, Jordan MI (2003) Latent dirichlet allocation. J Mach Learn Res 3(Jan):993–1022
- Braaksma JP, Cook WJ (1980) Human orientation in transportation terminals. Transp Eng J Am Soc Civil Eng 106(2):189–203
- Brunec IK, Nantais MM, Sutton JE, Epstein RA, Newcombe NS (2023) Exploration patterns shape cognitive map learning. Cognition 233:105360. https://doi.org/10.1016/j.cognition.2022.105360
- Bunk S, Krestel R (2018) WELDA: enhancing topic models by incorporating local word context. Paper presented at the Proceedings of the 18th ACM/IEEE on Joint Conference on Digital Libraries, Fort Worth, Texas, USA
- Büyükşahin S (2023) Effects of COVID-19 pandemic on spatial preferences and usage habits of users in shopping malls and its relation with circulation layout. Ain Shams Eng J 14(2):101838. https://doi.org/10.1016/j.asej. 2022.101838
- Cai H, Spreckelmeyer K (2022) The continuous learning cycle: a multi-phase post-occupancy evaluation (POE) of decentralized nursing unit design. HERD: Health Environ Res Design J 15(2):134–148. https://doi.org/10. 1177/19375867211051657
- Can I, Heath T (2016) In-between spaces and social interaction: a morphological analysis of Izmir using space syntax. J Hous Built Environ 31(1):31–49
- Can Traunmüller I, İnce Keller İ, Şenol F (2023) Application of space syntax in neighbourhood park research: an investigation of multiple socio-spatial attributes of park use. Local Environ 28(4):529–546. https://doi.org/10. 1080/13549839.2022.2160973
- Chiang Y-C, Li D (2019) Metric or topological proximity? The associations among proximity to parks, the frequency of residents' visits to parks, and perceived stress. Urban For Urban Green 38:205–214. https://doi. org/10.1016/j.ufug.2018.12.011
- Dahal B, Kumar SA, Li Z (2019) Topic modeling and sentiment analysis of global climate change tweets. Soc Netw Anal Min 9:1–20
- Dalton N (2001) Fractional configurational analysis and a solution to the Manhattan problem. In: Proceedings of the 3rd International Symposium on Space Syntax, vol 26. Georgia institute of Technology, Atlanta, Georgia, pp 1–13

- di Bella E, Corsi M, Leporatti L, Persico L (2015) The spatial configuration of urban crime environments and statistical modeling. Environ Plann B Urban Anal City Sci 44(4):647–667. https://doi.org/10.1177/0265813515 624686
- Domènech A, Gutiérrez A, Anton Clavé S (2020) Built environment and urban cruise tourists' mobility. Ann Tour Res 81:102889. https://doi.org/10. 1016/j.annals.2020.102889
- Edwards MJ (2013) The configuration of built space at Pataraya and Wari provincial administration in Nasca. J Anthropol Archaeol 32(4):565–576. https://doi.org/10.1016/j.jaa.2013.09.004
- El-Darwish II (2022) Enhancing outdoor campus design by utilizing space syntax theory for social interaction locations. Ain Shams Eng J 13(1):101524. https://doi.org/10.1016/j.asej.2021.06.010
- El-Hadedy N, El-Husseiny M (2022) Evidence-based design for workplace violence prevention in emergency departments utilizing CPTED and space syntax analyses. HERD Health Environ Res Design J 15(1):333–352. https://doi.org/10.1177/19375867211042902
- Elshater A, Abusaada H, Afifi S (2019) What makes livable cities of today alike? Revisiting the criterion of singularity through two case studies. Cities 92:273–291. https://doi.org/10.1016/j.cities.2019.04.008
- Esposito D, Santoro S, Camarda D (2020) Agent-based analysis of urban spaces using space syntax and spatial cognition approaches: a case study in Bari, Italy. Sustainability 12(11):4625. https://doi.org/10.3390/su121 14625
- Fang D, Yang H, Gao B, Li X (2018) Discovering research topics from library electronic references using latent Dirichlet allocation. Library Hi Tech
- Feeman LC (1978) Centrality in social networks conceptual clarification. Soc Netw 1:215–239. https://doi.org/10.1016/0378-8733(78)90021-7
- Ferdous F, Moore KD (2015) Field observations into the environmental soul: spatial configuration and Social Life for people experiencing Dementia. Am J Alzheimer's Dis Other Dementias<sup>®</sup> 30(2):209–218. https://doi.org/ 10.1177/1533317514545378
- Fladd SG (2017) Social syntax: an approach to spatial modification through the reworking of space syntax for archaeological applications. J Anthropol Archaeol 47:127–138. https://doi.org/10.1016/j.jaa.2017.05.002
- Freksa C (2004) Spatial cognition an AI perspective. In: Proceedings of the 16th European Conference on Artificial Intelligence, pp 1122–1128
- Freksa C, Klippel A, Winter S (2007) A cognitive perspective on spatial context. In: Dagstuhl Seminar Proceedings, 2007. Schloss Dagstuhl-Leibniz-Zentrum fr Informatik
- Geng S, Chau H-W, Yan S, Zhang W, Zhang C (2021) Comparative analysis of hospital environments in Australia and China using the space syntax approach. Int J Build Pathol Adapt 39(3):525–546. https://doi.org/10. 1108/JBPA-04-2020-0031
- Gibson JJ (1950) The perception of the visual world. Riverside Press, Cambridge
- Gümüş İ, Yılmaz E (2022) An evaluation of the relationship between everyday life rhythms and urban morphology: the square of Bursa Kent Meydani Shopping Center, Turkey. SAGE Open 12(2):21582440221104090. https://doi.org/10.1177/21582440221104090
- Güngör O, Harman Aslan E (2020) Defining urban design strategies: an analysis of Iskenderun city center's imageability. Open House Int 45(4):407–425. https://doi.org/10.1108/OHI-07-2020-0104
- Hajrasouliha A, Yin L (2014) The impact of street network connectivity on pedestrian volume. Urban Stud 52(13):2483–2497. https://doi.org/10. 1177/0042098014544763
- Hanson J (1998) Decoding homes and houses. Cambridge University Press, Cambridge
- Haq S, Luo Y (2012) Space syntax in healthcare facilities research: a review. HERD Health Environ Res Design J 5(4):98–117. https://doi.org/10.1177/ 193758671200500409
- Hayward SC, Franklin SS (1974) Perceived openness-enclosure of architectural space. Environ Behav 6(1):37–52
- Hegazi YS, Tahoon D, Abdel-Fattah NA, El-Alfi MF (2022) Socio-spatial vulnerability assessment of heritage buildings through using space syntax. Heliyon 8(3):e09133. https://doi.org/10.1016/j.heliyon.2022.e09133

- Heimerl F, Lohmann S, Lange S, Ertl T (2014) Word cloud explorer: text analytics based on word clouds. In: 2014 47th Hawaii International Conference on System Sciences, Waikoloa, HI, USA, 6–9 Jan. 2014. IEEE, IEEE Computer Society, CPS, pp 1833–1842. https://doi.org/10.1109/HICSS. 2014.231
- Hidayati I, Yamu C, Tan W (2021) Realised pedestrian accessibility of an informal settlement in Jakarta, Indonesia. J Urbanism: Int Res Placemaking Urban Sustain 14(4):434–456. https://doi.org/10.1080/17549175.2020.1814391
- Hillier B (1996) Space is the machine: a configurational theory of architecture. Cambridge University Press, Cambridge, MA
- Hillier B (1999) The hidden geometry of deformed grids: or, why space syntax works, when it looks as though it shouldn't. Environ Plan 26(2):169–191. https://doi.org/10.1068/b4125
- Hillier B (2012) Studying cities to learn about minds: some possible implications of space syntax for spatial cognition. Environ Plan 39(1):12–32. https://doi.org/10.1068/b34047t
- Hillier B, Hanson J (1984) The social logic of space. Cambridge University Press, Cambridge, MA
- Hillier B, lida S (2005) Network and psychological effects in urban movement. In: Cohn AG, Mark DM (eds) Spatial information theory. Springer, Berlin Heidelberg, pp 475–490
- Hillier B, Vaughan L (2007) The city as one thing. Prog Plan 67(3):205–230
- Hillier B, Hanson J, Graham H (1987) Ideas are in things: an application of the space syntax method to discovering house genotypes. Environ Plan 14(4):363–385
- Hillier B, Penn A, Hanson J, Grajewski T, Xu J (1993) Natural movement: or, configuration and attraction in Urban Pedestrian Movement. Environ Plan 20(1):29–66. https://doi.org/10.1068/b200029
- Hillier W, Yang T, Turner A (2012) Normalising least angle choice in depthmapand how it opens up new perspectives on the global and local analysis of city space. J Space Syntax 3(2):155–193
- Hölscher C, Brösamle M, Vrachliotis G (2012) Challenges in multilevel wayfinding: a case study with the space syntax technique. Environ Plan 39(1):63–82. https://doi.org/10.1068/b34050t
- Istiani NFF, Alkadri MF, van Nes A, Susanto D (2023) Investigating the spatial network of playgrounds during covid-19 based on a space syntax analysis case study: 10 playgrounds in Delft, the Netherlands. Cogent Soc Sci 9(1):2163754. https://doi.org/10.1080/23311886.2022.2163754
- Jayasinghe A, Sano K, Rattanaporn K (2017) Application for developing countries: estimating trip attraction in urban zones based on centrality. J Traffic Transp Eng (English Edition) 4(5):464–476. https://doi.org/10. 1016/j.jtte.2017.05.011
- Jelodar H, Wang Y, Yuan C, Feng X, Jiang X, Li Y, Zhao L (2019) Latent dirichlet allocation (LDA) and topic modeling: models, applications, a survey. Multimedia Tools Appl 78(11):15169–15211. https://doi.org/10.1007/ s11042-018-6894-4
- Jiang B (2009) Ranking spaces for predicting human movement in an urban environment. Int J Geogr Inf Sci 23(7):823–837. https://doi.org/10.1080/ 13658810802022822
- Jiang B, Claramunt C (2004) Topological analysis of urban street networks. Environ Plan 31(1):151–162. https://doi.org/10.1068/b306
- Jiang B, Jia T (2011) Agent-based simulation of human movement shaped by the underlying street structure. Int J Geogr Inf Sci 25(1):51–64. https:// doi.org/10.1080/13658811003712864
- Jiang B, Liu C (2009) Street-based topological representations and analyses for predicting traffic flow in GIS. Int J Geogr Inf Sci 23(9):1119–1137. https://doi.org/10.1080/13658810701690448
- Joshi R, Ossmann M, Joseph A (2023) Measuring potential visual exposure of physicians during shift-end handoffs and its impact on interruptions, privacy, and collaboration. HERD Health Environ Res Design J 16(1):175–199. https://doi.org/10.1177/19375867221131934
- Kaplan R, Kaplan S (1989) The experience of nature. A psychological perspective. Cambridge University Press, Cambridge
- Karthika PS, Kedar V, Verma A (2022) A walk accessibility-based approach to assess crowd management in mass religious gatherings. J Transp Geogr 104:103443. https://doi.org/10.1016/j.jtrangeo.2022.103443
- Katarya R, Das I, Shrivastava B, Keswani K (2022) An evaluation of LDA modeling in financial news articles. In: Sanyal G, Travieso-González CM, Awasthi S, Pinto CMA, Purushothama BR (eds) International Conference on Artificial Intelligence and Sustainable Engineering, Singapore, Springer Nature Singapore, pp 365–379

Kee YH, Li C, Kong LC, Tang CJ, Chuang K-L (2019) Scoping review of mindfulness research: a topic modelling approach. Mindfulness 10(8):1474– 1488. https://doi.org/10.1007/s12671-019-01136-4

Keszei B, Halász B, Losonczi A, Dúll A (2019) Space syntax's relation to seating choices from an evolutionary approach. Periodica Polytech Archit 50(2):115–123. https://doi.org/10.3311/PPar.14251

Khotbehsara EM, Askarizad R, Mehrinejad M, Nasab SN, Somasundaraswaran K (2023) The impact of COVID-19 on visitors' wayfinding within healthcare centers. Ain Shams Eng J 14(5):101957. https://doi.org/10.1016/j.asej. 2022.101957

Kim JY, Kim YO (2023) Analysis of pedestrian behaviors in subway station using agent-based model: case of Gangnam Station, Seoul, Korea. Buildings 13(2):537

Kim YO, Penn A (2004) Linking the spatial syntax of cognitive maps to the spatial syntax of the environment. Environ Behav 36(4):483–504

Kim M, Piao G (2017) A study on the applying concept of metric weighting to space syntax. J Asian Archit Build Eng 16(3):447–454. https://doi.org/10. 3130/jaabe.16.447

Knöll M, Neuheuser K, Cleff T, Rudolph-Cleff A (2018) A tool to predict perceived urban stress in open public spaces. Environ Plann B Urban Anal City Sci 45(4):797–813. https://doi.org/10.1177/0265813516686971

Koohsari MJ, Owen N, Cerin E, Giles-Corti B, Sugiyama T (2016) Walkability and walking for transport: characterizing the built environment using space syntax. Int J Behav Nutr Phys Act 13(1):121. https://doi.org/10.1186/ s12966-016-0448-9

Lee JH, Ostwald MJ (2020) Grammatical and syntactical approaches in architecture: emerging research and opportunities. IGI Global, Hershey, PA

Lee JH, Ostwald MJ, Gu N (2016) A justified Plan Graph (JPG) grammar approach to identifying spatial design patterns in an architectural style. Environ Plann B Urban Anal City Sci 45(1):67–89. https://doi.org/10. 1177/0265813516665618

Lee JH, Ostwald MJ, Lee H (2017) Measuring the spatial and social characteristics of the architectural plans of aged care facilities. Front Architectural Res 6(4):431–441. https://doi.org/10.1016/j.foar.2017.09.003

Lee JH, Ostwald MJ, Dawes MJ (2022) Examining visitor-inhabitant relations in Palladian Villas. Nexus Netw J 24(2):315–332. https://doi.org/10.1007/ s00004-021-00589-2

- Lee JH, Ostwald MJ, Zhou L (2023) Socio-spatial experience in Space syntax research: a PRISMA-compliant review. Buildings 13(3):644
- Li R, Klippel A (2012) Wayfinding in libraries: can problems be predicted? J Map Geogr Libr 8(1):21–38. https://doi.org/10.1080/15420353.2011.622456

Li J, Shu Y, Chen N, Wang F, Li H (2022) Re-socialisation'in isolated spaces: a case study on the social organisation of Fangcang shelter hospital patients under extreme spatial conditions. Indoor Built Environ 31(5):1210–1223. https://doi.org/10.1177/1420326x20973745

Liu P, Xiao X, Zhang J, Wu R, Zhang H (2018) Spatial configuration and online attention: a space syntax perspective. Sustainability 10(1):221. https:// doi.org/10.3390/su10010221

Ma D, Omer I, Osaragi T, Sandberg M, Jiang B (2018) Why topology matters in predicting human activities. Environ Plann B Urban Anal City Sci. https://doi.org/10.1177/2399808318792268

Mansouri M, Ujang N (2017) Space syntax analysis of tourists' movement patterns in the historical district of Kuala Lumpur, Malaysia. J Urban Int Res Placemaking Urban Sustain 10(2):163–180. https://doi.org/10.1080/ 17549175.2016.1213309

Marcus L, Koch D (2016) Cities as implements or facilities—the need for a spatial morphology in smart city systems. Environ Plann B: Urban Anal City Sci 44(2):204–226. https://doi.org/10.1177/0265813516685565

Marcus L, Giusti M, Barthel S (2016) Cognitive affordances in Sustainable Urbanism: contributions of space syntax and spatial cognition. J Urban Des 21(4):439–452

Marquardt G, Johnston D, Black BS, Morrison A, Rosenblatt A, Lyketsos CG, Samus QM (2011) Association of the spatial layout of the home and ADL abilities among older adults with Dementia. Am J Alzheimer's Dis Other Dementias<sup>®</sup> 26(1):51–57. https://doi.org/10.1177/1533317510 387584

Mimno D, Wallach H, Talley E, Leenders M, McCallum A (2011) Optimizing semantic coherence in topic models. In: Proceedings of the 2011 conference on empirical methods in natural language processing, pp 262–272 Mohamed AA, Stanek D (2020) The influence of street network configuration on sexual harassment patterns in Cairo. Cities 98:102583. https://doi. org/10.1016/j.cities.2019.102583

Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 6(7):e1000097. https://doi.org/10.1371/journal.pmed. 1000097

Morales J, Flacke J, Zevenbergen J (2017) Modelling residential land values using geographic and geometric accessibility in Guatemala City. Environ Plann B: Urban Anal City Sci. https://doi.org/10.1177/23998 08317726332

Neo JRJ, Sagha-Zadeh R (2017) The influence of spatial configuration on the frequency of use of hand sanitizing stations in health care environments. Am J Infect Control 45(6):615–619. https://doi.org/10.1016/j. ajic.2017.01.033

Netto VM (2016) What is space syntax not? Reflections on space syntax as sociospatial theory. Urban Des Int 21(1):25–40. https://doi.org/10. 1057/udi.2015.21

Nicoletta S, Eletta N, Cardinali P, Migliorini L (2022) A broad study to develop maternity units design knowledge combining spatial analysis and mothers' and midwives' perception of the birth environment. HERD Health Environ Res Design J 15(4):204–232. https://doi.org/10.1177/ 19375867221098987

Nubani L, Puryear A, Kellom K (2018) Measuring the effect of visual exposure and saliency of museum exhibits on visitors' level of contact and engagement. Behav Sci 8(11):100. https://doi.org/10.3390/bs8110100

O'Hara S, Klar RT, Patterson ES, Morris NS, Ascenzi J, Fackler JC, Perry DJ (2018) Macrocognition in the Healthcare built Environment (mHCBE): a focused Ethnographic Study of neighborhoods in a Pediatric Intensive Care Unit. HERD Health Environ Res Des J 11(2):104–123. https:// doi.org/10.1177/1937586717728484

Omer I, Goldblatt R (2017) Using space syntax and Q-analysis for investigating movement patterns in buildings: the case of shopping malls. Environ Plann B: Urban Anal City Sci 44(3):504–530. https://doi.org/ 10.1177/0265813516647061

Omer I, Jiang B (2015) Can cognitive inferences be made from aggregate traffic flow data? Computers. Environ Urban Syst 54:219–229. https:// doi.org/10.1016/j.compenvurbsys.2015.08.005

Ostwald MJ (2011) A justified plan graph analysis of the early Houses (1975–1982) of Glenn Murcutt. Nexus Netw J 13(3):737–762

Ozbil A, Yesiltepe D, Argin G, Rybarczyk G (2021) Children's active school travel: examining the combined perceived and objective built-environment factors from space syntax. Int J Environ Res Public Health 18(1):286. https://doi.org/10.3390/ijerph18010286

Pafka E, Dovey K, Aschwanden GDPA (2020) Limits of space syntax for urban design: axiality, scale and sinuosity. Environ Plann B Urban Anal City Sci 47(3):508–522. https://doi.org/10.1177/2399808318786512

Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 372:n71. https://doi.org/10.1136/bmj.n71

Pagkratidou M, Galati A, Avraamides MN (2020) Do environmental characteristics predict spatial memory about unfamiliar environments? Spat Cogn Comput 20:1–32

Pan M, Shen Y, Jiang Q, Zhou Q, Li Y (2022) Reshaping publicness: research on correlation between public participation and spatial form in urban space based on space syntax—a case study on Nanjing Xinjiekou. Buildings 12(9):1492

Penn A (2003) Space syntax and spatial cognition: or why the Axial line? Environ Behav 35(1):30–65. https://doi.org/10.1177/0013916502 238864

Raban J (1974) Soft City. Hamish Hamilton, London

Rashid M, Boyle DK, Crosser M (2014) Network of spaces and Interactionrelated behaviors in adult intensive care units. Behav Sci 4(4):487–510. https://doi.org/10.3390/bs4040487

Rashid M, Khan N, Jones B (2016) Physical and visual accessibilities in intensive care units: a comparative study of open-plan and racetrack units. Crit Care Nurs Q 39(4):313–334. https://doi.org/10.1097/CNQ.0000000000000000126

- Ratti C (2004) Space syntax: some inconsistencies. Environ Plan 31(4):487–499. https://doi.org/10.1068/b3019
- Röder M, Both A, Hinneburg A (2015) Exploring the space of topic coherence measures. In: Proceedings of the eighth ACM international conference on Web search and data mining, pp 399–408
- Safaie Ghamsary E, Karimimoshaver M, Akhavan A, Afzali Goruh Z, Aram F, Mosavi A (2023) Locating pocket parks: assessing the effects of land use and accessibility on the public presence. Environ Sustain Indic 18:100253. https://doi.org/10.1016/j.indic.2023.100253
- Scoppa MD, Peponis J (2015) Distributed attraction: the effects of street network connectivity upon the distribution of retail frontage in the City of Buenos Aires. Environ Plan 42(2):354–378. https://doi.org/10.1068/ b130051p
- Serra M, Hillier B (2019) Angular and metric distance in road network analysis: a nationwide correlation study. Comput Environ Urban Syst 74:194– 207. https://doi.org/10.1016/j.compenvurbsys.2018.11.003
- Sharmin S, Kamruzzaman M (2018) Meta-analysis of the relationships between space syntax measures and pedestrian movement. Transp Rev 38(4):524–550. https://doi.org/10.1080/01441647.2017.1365101
- Shatu F, Yigitcanlar T, Bunker J (2019) Shortest path distance vs. least directional change: empirical testing of space syntax and geographic theories concerning pedestrian route choice behaviour. J Transp Geogr 74:37–52. https://doi.org/10.1016/j.jtrangeo.2018.11.005
- Shen Y, Wu Z (2022) Functional visibility graph analysis: quantifying visuofunctional space with social media check-in data. Environ Plann B Urban Anal City Sci 49(1):41–57. https://doi.org/10.1177/23998083211001840
- Sheng Q, Wan D, Yu B (2021) Effect of space configurational attributes on social interactions in urban parks. Sustainability 13(14):7805. https://doi.org/10.3390/su13147805
- Sohn D-W (2016) Residential crimes and neighbourhood built environment: assessing the effectiveness of crime prevention through environmental design (CPTED). Cities 52:86–93. https://doi.org/10.1016/j.cities.2015. 11.023
- Srinurak N, Mishima N (2017) Urban Axis and City shape evaluation through spatial configuration in 'Lan na' Northern Thailand historic city. City Terri Archit 4(1):10. https://doi.org/10.1186/s40410-017-0067-z
- Stamps AE (2005) Isovists, enclosure, and permeability theory. Environ Plan 32(5):735–762. https://doi.org/10.1068/b31138
- Steadman PJ (1983) Architectural morphology. Pion, London
- Stevens K, Kegelmeyer P, Andrzejewski D (2012) Buttler D Exploring topic coherence over many models and many topics. In: Proceedings of the 2012 joint conference on empirical methods in natural language processing and computational natural language learning, pp 952–961
- Stracqualursi L, Agati P (2022) Tweet topics and sentiments relating to distance learning among Italian Twitter users. Sci Rep 12(1):9163. https://doi.org/ 10.1038/s41598-022-12915-w
- Syed S, Spruit M et al (2017) Full-text or abstract? examining topic coherence scores using latent dirichllocation. In: 2017 IEEE International conference on data science and advanced analytics (DSAA) IEEE, pp 165–174
- Turner A (2001) Depthmap: a program to perform visibility graph analysis. In: Proceedings 3rd International Symposium on Space Syntax. Georgia Institute of Technology, Atlanta, GA, pp 31.31–31.39
- Turner A (2007) From Axial to Road-Centre lines: a new representation for space syntax and a New Model of Route Choice for Transport Network Analysis. Environ Plan 34(3):539–555. https://doi.org/10.1068/b32067
- Turner A, Doxa M, O'Sullivan D, Penn A (2001) From isovists to visibility graphs: a methodology for the analysis of architectural space. Environ Plan 28(1):103–121. https://doi.org/10.1068/b2684
- Tzeng S-Y, Huang J-S (2009) Spatial forms and signage in Wayfinding decision points for hospital outpatient services. J Asian Archit Build Eng 8(2):453–460. https://doi.org/10.3130/jaabe.8.453
- Vaughan L (2015) Glossary of Space Syntax. In: Vaughan L (ed) Suburban urbanities: suburbs and the life of the High Street. UCL Press, London, UK, pp 307–312
- Wang L, Han X, He J, Jung T (2022) Measuring residents' perceptions of city streets to inform better street planning through deep learning and space syntax. ISPRS J Photogrammetry Remote Sens 190:215–230. https://doi.org/10.1016/j.isprsjprs.2022.06.011

- Wu Y, Liu M, Zheng WJ, Zhao Z, Xu H (2012) Ranking gene–drug relationships in biomedical literature using latent dirichlet allocation. Biocomputing 2012. World Scientific, pp 422–433
- Xu Y, Chen X (2022) The spatial vitality and spatial environments of urban underground space (UUS) in metro area based on the spatiotemporal analysis. Tunn Undergr Space Technol 123:104401. https://doi.org/10. 1016/j.tust.2022.104401
- Yang Y, Vaughan L (2022) Does area type matter for pedestrian distribution? Testing movement economy theory on gated and non-gated housing estates in Wuhan, China. Computers Environ Urban Syst 97:101868. https://doi.org/10.1016/j.compenvurbsys.2022.101868
- Yang C, Qian Z, Li H (2023) Measuring the accessibility deprivation of concentrated resettlement communities in China: an integrated approach of space syntax and multi-criteria decision analysis. Appl Geogr 150:102839. https://doi.org/10.1016/j.apgeog.2022.102839
- Yaseen FR, Mustafa FA (2023) Visibility of nature-connectedness in school buildings: an analytical study using biophilic parameters, space syntax, and space/nature syntax. Ain Shams Eng J 14(5):101973. https://doi. org/10.1016/j.asej.2022.101973
- Yıldırım ÖC, Çelik E (2023) Understanding pedestrian behavior and spatial relations: a pedestrianized area in Besiktas, Istanbul. Front Archit Res 12(1):67–84. https://doi.org/10.1016/j.foar.2022.06.009
- Zeng M, Wang F, Xiang S, Lin B, Gao C, Li J (2020) Inheritance or variation? Spatial regeneration and acculturation via implantation of cultural and creative industries in Beijing's traditional compounds. Habitat Int 95:102071. https://doi.org/10.1016/j.habitatint.2019.102071
- Zerouati W, Bellal T (2020) Evaluating the impact of mass housings' in-between spaces' spatial configuration on users' social interaction. Front Architect Res 9(1):34–53. https://doi.org/10.1016/j.foar.2019.05.005
- Zhai Y, Baran PK (2016) Do configurational attributes matter in context of urban parks? Park pathway configurational attributes and senior walking. Landsc Urban Plann 148:188–202. https://doi.org/10.1016/j.landu rbplan.2015.12.010
- Zhai Y, Korça Baran P, Wu C (2018) Can trail spatial attributes predict trail use level in urban forest park? An examination integrating GPS data and space syntax theory. Urban For Urban Green 29:171–182. https://doi. org/10.1016/j.ufug.2017.10.008
- Zhang L, Chiaradia AJ (2022) Walking in the cities without ground, how 3d complex network volumetrics improve analysis. Environ Plann B Urban Anal City Sci 49(7):1857–1874. https://doi.org/10.1177/2399808321 1070567
- Zhang T, Lian Z, Xu Y (2020) Combining GPS and space syntax analysis to improve understanding of visitor temporal–spatial behaviour: a case study of the Lion Grove in China. Landsc Res 45(4):534–546. https://doi. org/10.1080/01426397.2020.1730775
- Zhao X, Hong K (2023) Basic analysis of the correlation between the accessibility and utilization activation of public libraries in seoul: focusing on location and subway factors. Buildings 13(3):600
- Zheng J, Bai X, Wu Z, Zhang S, Zhang T, Wang H (2022) Research on the spatial behavior conflict in suburban village communities based on GPS tracking and cognitive mapping. J Asian Archit Build Eng 21(6):2605–2620. https://doi.org/10.1080/13467581.2021.1971680

#### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.