Landscape Perception through Complex Data: Exploring George Hargreaves's Queen Elizabeth Olympic Park in London

Paola Riva[1], Paola Sturla[2], Anna Calissano[1], Simone Vantini[1]

[1]Politecnico di Milano, Milan/Italy
[2]Harvard Graduate School of Design, Massachusetts/USA

Abstract: This paper discusses a case study of users' perceptions of the Queen Elizabeth Olympic Park in London through an object-oriented data analysis of the text of Tweets. Starting with the gathering of text data from Twitter, the interdisciplinary team of researchers explores the potential of semantic analysis to provide designers with a tool for describing the experience of a place.

Keywords: Object-oriented data analysis, statistical text analysis, text embedding, perception, Queen Elizabeth Olympic Park, exploration, interdisciplinary studies

1 Introduction

Over the past decade, the pervasive presence of digital mobile devices together with the spread of sensors across urban spaces has generated a mass of data, often referred to as "Big Data" (MANYIKA et al. 2011), which are usually studied through algorithmic automated analytical tools, as for example machine learning (SAMUEL 1959). The academic literature has discussed Big Data characteristics and epistemology extensively (BOYD & CRAWFORD 2016, KITCHIN 2014B, KITCHIN & DODGE 2011).

These studies have contributed to a productive debate on the nature of such information, laying the foundation for experiments on potential applications in the field of the complexity theory of cities and smart cities (BATTY 2005, 2013, PORTUGALI & HAKEN 2000, 2011, PORTUGALI et al. 2012). However, from a disciplinary perspective, the interesting question is: What are these analyses for?

2 A Landscape Design Disciplinary Perspective

Within the design disciplines, landscape design studies the relationship between the subject and the environment through the means of design. The physical features of outdoor spaces and their form, materiality, and evolution through time shape their "sensible" qualities (KANT 1900 [1781], 1992 [1791]), generating an immersive environment. The use of the word "environment" suggests that such relationships include the immaterial ones. Environment refers to the concept of *Umwelt,* defined in 1926 by Jakob von Uexküll, a German biologist. *Umwelt* can be translated into English as "environment". However, it differs from the mere context because of the active role of the subject in perceiving it.

The primary assumption of this research is that landscapes are complex systems, where emergent characteristics can be found and described on at least two levels, as discussed by scholars in the complexity science field, such as Gregory Bateson (BATESON 1972). At first, charac-

Journal of Digital Landscape Architecture, 4-2019, pp. 42-56. © Wichmann Verlag, VDE VERLAG GMBH · Berlin · Offenbach. ISBN 978-3-87907-663-5, ISSN 2367-4253, e-ISSN 2511-624X, doi:10.14627/537663005.

teristics emerge in the interaction of physical objects with environmental systems and phenomena over time. Secondly, they emerge in the aesthetic experience of the landscape itself. According to JORGEN DEHS (2001), "landscape means a flowing variety that is held together by the spectator's glance", suggesting that the "sense" of a landscape can be conceived through a concept of "identity in translation", the relationship between its physical characteristics and the unique, individual perception of it.

There is not a unique and shared definition of landscape[1]. For the sake of this research, the focus is on the user's aesthetic experience. The authors borrow the definition of aesthetic experience from Immanuel Kant's work, meaning the pleasure associated with the moment of judging something beautiful. To impress the user in this regard may be one of the goals of a landscape project, beyond its functional, environmental, and standards-compliancy requirements.

3 Big Data and Landscape Design: What For?

The interest in Big Data in the field of landscape design has been increasing over recent years. Papers highlight the potential for Big Data analysis in users' landscape perceptions (KAUSSEN 2018, MONTANO 2018, ROTH et al. 2018), human experiences of places (TULLOCH & IM 2018), and the relationship between public spaces and behaviour (JIE & YUNING 2018). Within the context of the planning disciplines, data analysis is commonly used to quantitatively estimate flows and movements. For example, over the past decade traffic studies have been carried out that take advantage of geo-localized spatial data from smartphones for both research and implementation of "smart city" applications. The introduction of behavioural studies (CRANENBURGH, ALWOSHEEL, N.D. & KWAN et al. 2007) highlights emergent decision-making patterns if combined with the analysis of people's movements in urban spaces (KITCHIN 2014A, MANFREDINI ET AL. 2016, RATTI et al. 2006), suggesting innovative approaches in policy design, implementation, and evaluation (BRUIJN & HEUVELHOF 2008, NIKAYIN & REUVER 2012, PUCCI 2016), as stated also by PoliMi POLIVISU researchers[2]. The continuous stress on quantitative research has pushed toward a positivistic approach to Big Data analysis in the design field. In more recent years, research is shifting toward a reconsideration of the aesthetic values of landscapes through a rediscovery of cybernetics and complexity science by mid-XX[th] century authors. According to KAREN M'CLOSELY and KEITH VANDERSYS (2017) "designers are increasingly asked to substantiate their work through metrics. However, these criteria alone do not encompass the full value of landscapes. The pattern recognition that Bateson […] argued for […] is first and foremost an aesthetic framework".

The reason that the positivistic approach to data in planning disciplines is problematic is easy to grasp if one goes back to HERBERT SIMON'S definition of the artifact and the artificial: "But you will have to understand me as using artificial in as neutral a sense as possible: as meaning man-made, as opposed to natural" and "the world we live in today is much more a

---

1 For a broader definition of landscape see JOHN STILGOE (2015), What Is Landscape? and MICHAEL JAKOB (2018), What Is Landscape?

2 https://www.polivisu.eu

man-made, or artificial, world than it is a natural world. Almost every element in our environment show's evidence of men's artifice" (1969). Following his argument, the artifact is, therefore, the interface between the inner environment (the object) and the outer environment, or the context. As a consequence, complex systems such as cities or landscapes are facts of art, mediated by humans (both designers and users). Therefore, the mere application of natural science paradigms cannot lead to their description. JUVAL PORTUGALI, who researches the complexity theory of cities (CTC), stated in 2012 that the missing link between CTC and cognition might lead to "irrelevant urban studies?" (PORTUGALI et al. 2012).

How does one develop the missing link? Could Big Data and their analysis provide us with a glance at users' perceptions of a landscape? Through Big Data can one identify with the person seated on the bench, or the one walking through space, and experience the landscape through his/her eyes?

These elementary questions are the basis for the experiment described in this article. From the landscape design disciplinary perspective, the research aims to narrate the users' experience of a high-end landscape designed by Georges Hargreaves, the Queen Elizabeth Olympic Park (QEOP) in London (Figures 1 and 2), through an object-oriented data analysis (MARRON & ALONSO 2014, WANG & MARRON 2007) of visitors' Tweets published in the area. The case study has been selected for three main reasons: first, the site is located in an English-speaking city. Second, the landscape design has a framework of transit-oriented development (TOD), implying an intention to design connective spaces. Finally, the capacity of the stadium allows for big events, enabling a wider gathering of Twitter data. The goal is to observe whether the landscape, the park, and its recognizable design are among the most common discussion topics in the area. It is important to note that the output of the analysis is not meant to provide any measurable metric to the designer and that we cannot infer a judgment on the quality of the landscape from the output. It is rather an impression, a sort of painting that narrates an experience, a vision of the landscape interpreted by the users and recorded thanks to Urbanscope (ARNABOLDI et al. 2017), a combination of the analytical tools and the researchers' domain knowledge. It delivers qualitative information and constitutes a new point of view on the landscape. To capture users' impressions concerning the park, a research approach inspired by the concept of Macroscope (ROSNAY 1979) was utilized, following the process framework introduced by ARNABOLDI et al. Macroscope is a holistic approach to studying the complexity of a system, mixing competencies and views of different field experts. The polyhedric research team is the Macroscope, i. e., a new "tool" to address complexity using a multidisciplinary approach.

Fig. 1: Hargreaves Associates, Queen Elizabeth Olympic Park, London, UK. The park during an Olympic event (Image © Hargreaves Associates)

Fig. 2: Hargreaves Associates, Queen Elizabeth Olympic Park, London, UK. The park on a regular weekday (Image © Hargreaves Associates)

4 Text as a Complex Object

As previously stated, the analysis aims to capture visitors' perceptions of the Queen Elizabeth Olympic Park. The impression can be expressed via different media, such as photos, videos, images, and texts. In this experiment, we focus on visitors' textual impressions. Several different sources of textual feedback are available: newspaper articles, blog posts, personal diaries, etc. Among these, we decided to analyse textual impressions shared on social media. The main reason was that social media posts could be geo-localized (if position access was allowed by the user), offering a unique chance to track opinions in space. Among the existing social media platforms, we focused on Twitter. This platform was selected for two main reasons. First, Twitter has a public philosophy, with user profiles open by default and the possibility to freely download a subsample of Twitter data (CALISSANO et al. 2018). Second, it is a micro-blogging platform, in which users share in less than 280 characters their own opinion about specific topics. The sharpness of the posts helps the analysis of the impression because users are forced to be concise (CALISSANO et al. 2018). Data were downloaded from Twitter during two weeks of March 2018, filtering Tweets according to the geographic location associated with the post.

Even though the development of quantitative techniques to analyse text data dates back to the early 1980s (HOBBS et al. 1982), the rapidly growing amount of user-generated contents in the digital era stresses the requirement for high-performing and highly interpretable techniques to understand human language. This resulted in development of a wide variety of methods to analyse the text. From a statistical perspective, very few studies tackle the issue of defining a suitable mathematical framework to embed and analyse the complex properties of human language (RIVA 2018). To fill this gap, this paper proposes a statistically innovative way to approach the analysis of textual data, which are perfect examples of complex data (RIVA 2018). Indeed, statistically speaking, a complex datum is a datum that cannot be naturally embedded in a Euclidean space (VANTINI 2018, WANG & MARRON 2007). This leads to a series of open questions regarding the analytic procedure. How can we measure the distance between the contents of two texts? How can we affirm that one text is similar to another? With the aim of approaching these issues, this paper offers a new viewpoint on the problem of embedding complex properties of language for statistical analysis, creating a novel connection between the complexities of text and the statistical framework of object-oriented data analysis. With the goal of understanding the impressions of visitors within a specific area, this paper demonstrates a novel application of object-oriented data analysis techniques to the text of Tweets, aimed at embedding Tweets in a mathematically "comfortable" space.

Once we collected a set of Tweets, the intuitive question that arose was whether or not all these texts treated similar topics. To address this question, we decided to perform the analysis in two distinct steps. First, we looked for a meaningful and manageable mathematic space where text could be embedded. This space needed to ensure the possibility to both analyse the text once it had been embedded and provide a meaningful interpretation of the embedding. In the second step, these embedded texts were analysed. In particular, we searched for differences among Tweets in an unsupervised fashion, performing cluster analysis. Without supposing any a priori difference among texts, we grouped Tweets according to their similarities in the embedding space, finding four different clusters of texts, i. e., four different visitors' impressions. During the specific time interval between 16 March and 3 April 2018, we observed visitors to the QEOP discussing on Twitter two main London Stadium events (i. e., football and rugby matches), visitors interested in the Westfield Mall (i. e., shopping-related Tweets), and visitors sharing impressions about the Queen Elizabeth Olympic Park. Observe that even if the groups of Tweets were not created by dividing contents according to their geographic location, the results of the analyses showed correspondence between the group a Tweet belonged to and its location.

4.1 Tweets' Collection and Sample Quality

As part of its public philosophy, Twitter allows users to access and download Tweets' information through Twitter API (Twitter Inc., Developer Documentation 2018). The platform offers a variety of connections, which differ in costs, sampling strategies, access level to specific contents, and time windows to download Tweets (Calissano et al. 2018, Twitter Inc., Developer Documentation 2018). With the goal of studying perceptions of visitors of the QEOP, the location of Tweets was selected as a key characteristic for collection from posts in order to ensure the analyses of thoughts of *actual* visitors to the park. The requirement for geographic information from Tweets, combined with the possibility of accessing up to 1% of all Tweets available worldwide establishing a real-time and free-of-charge connection with the platform, led to the choice of Twitter Streaming API with geographic parameters as the sampling strategy for the Tweets' collection (Twitter Inc., Twitter Streaming API 2018). The streaming connection established with Twitter from 6 PM of 16 March 2018 to 3 AM of 3 April 2018 resulted in the collection of 208,085 Tweets containing geographic information. These Tweets were collected by defining as streaming parameters the following two bounding boxes (Figure 3):

area A1: in between 0°01'90''W and 0°01'23''W, 51°54'04''N and 51°53'26''N

area A2: in between 0°01'87''W and 0°00'18''W, 51°54'52''N and 51°53'71''N

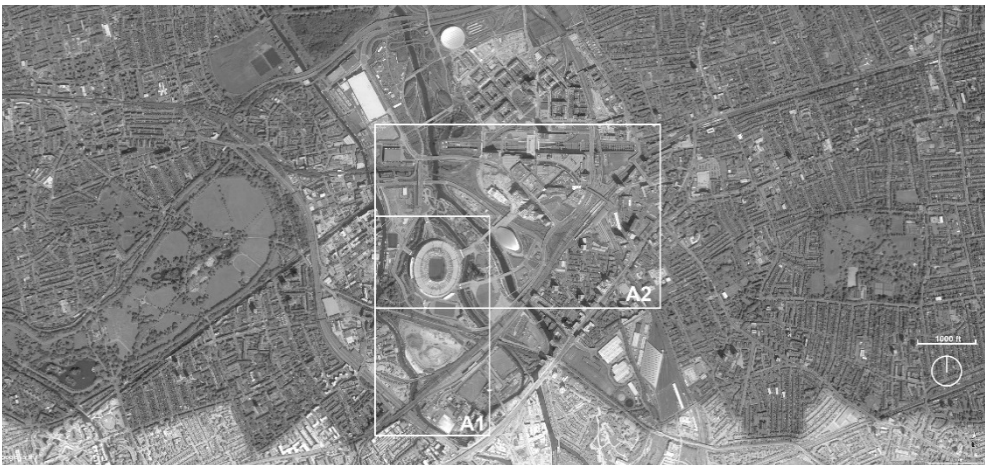

Fig. 3: Twitter data gathering areas on Google Earth aerial image

The position of collected Tweets was directly connected to user's location thanks to Twitter Places[3], a system used since 2010 by Twitter to assign geographic positions according to the user's desire to share her/his own GPS location or tagging a so-called "Twitter Place", a place represented on Twitter through a bounding box (TWITTER INC., GEO OBJECTS 2018). The Twitter Places system allows the definition of geographic regions with various extensions, from a single point on the Earth to an entire country (TWITTER INC. 2018). As described in Twitter documentation, Twitter Place logic combined with the use of Twitter Streaming API with geographic parameters resulted in gathering Tweets for which either the GPS coordinates fell in the bounding boxes defined as collection parameters, or there was a Twitter Place tag whose description intersected the bounding boxes defined as streaming parameters. The major consequence of this strategy, the use of the two alternatives for Tweet collection, was collection of Tweets that had not been published inside the regions defined as streaming parameters (Figure 4).

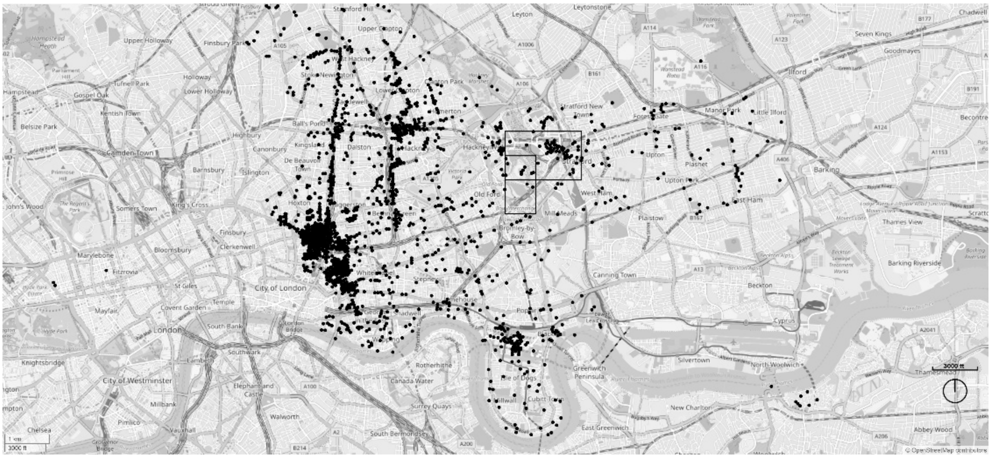

Fig. 4: Tweets gathered using geographic filters defined by given areas on Open Street Map

For instance, since the geographic representation of the city of London on the Twitter platform was wider than the two bounding boxes of Pudding and Stratford defined for the current Tweet collection, Tweets adding a "London" tag were gathered during the collection process, independently of the geographic position declared by the Twitter user. To capture perceptions of *actual* visitors to the two regions of the Queen Elizabeth Olympic Park, we decided to consider only those Tweets actually published inside the two bounding boxes of interest and generated by English-speaking Twitter users. Therefore, starting with the 179, 180 posts written in English language, the geographic limitation just described led to the analysis of 279 Tweets, which corresponded to messages published directly on the Twitter platform by English-speaking Twitter users who visited the two regions of park between 16 March and 3 April 2018 and shared their geographic position or tagged a place inside the QEOP.

[3] https://blog.twitter.com/official/en_us/a/2010/twitter-places-more-context-for-your-tweets.html

4.2 The Model: Tweets’ Embedding and Analysis

The choice of studying the content of Tweets to offer insights into the perceptions of visitors to the QEOP results in the complicated task of statistical analysis of written text, which is an expression of a highly complex system, namely, the human language (RIVA 2018). Although very few studies tackle the issue of defining a suitable mathematical framework to embed and statistically analyse the complex properties of human language (RIVA 2018), this paper creates a novel connection between the intrinsic complexities of text and the statistical framework of object-oriented data analysis (OODA). With the aim of contributing to understanding of the impressions of visitors within the specific area of Queen Elizabeth Olympic Park, this paper describes a novel application of OODA techniques to the statistical analysis of the text of Tweets, embedding semantic properties of Tweets in a mathematically meaningful and manageable space.

One of the basic approaches to semantic representation of text is the one of vector space models (VSMs) (SALTON, G., WONG, A. & YANG, C. 1975A, SALTON, G., YANG, C. & YU, C. 1975B). Inside this class, the Bag-of-Words (BoW) model considers the semantic of a Tweet as a bag, filled with words extracted from the text of the Tweet, where each word has a specific weight given by the number of occurrences of the term in the set of texts. Mathematically speaking, the Bag-of-Words model represents a Tweet as a vector, where each entry corresponds to a specific word and the associated value expresses the weight of the term in defining the semantic of the Tweet. An important implication of this representation is that the mathematical space in which these Tweets lie is a vector space, where Tweets’ representations (i. e., vectors) that are close to each other correspond to messages that express semantically similar contents.

Thanks to this particular property, vector space models have been largely used, and thus enriched and extended to consider a variety of more complex linguistic problematics (RIVA 2018). Among the interesting extensions, the class of distributional semantic models (DSMs) approaches the problem of understanding the meaning of groups of words (*compositional semantics*) rather than studying the meaning of individual terms (*lexical semantics*). Based on FIRTH’s expression “You should know a word by the company it keeps” (1957), these models assume the distributional hypothesis (HARRIS 1954) according to which similar words appear in the same context.

Among DSM models, the latent semantic model, commonly known as latent semantic analysis (LSA) (DEERWESTER et al. 1990) or latent semantic indexing (LSI) (DEERWESTER et al. 1988), is here used to represent Tweets in a mathematically meaningful and manageable space, with enhanced semantic properties with respect to the one of the vector space models (RIVA 2018). Indeed, the idea behind latent semantic models is to elaborate the textual information inside the collection of documents, here Tweets, to determine a certain number of *latent semantic concepts* addressed by these messages. Each of these latent semantic concepts is a mix of semantically similar words which together define a common notion, generally referred to as a “concept”, whose meaning would not be immediately apparent, therefore “latent”, when looking at the whole set of available words in Tweets. Therefore, latent semantic models do not embed texts in the space of all possible words but in a subspace constituted of concepts rather than single words. LSA results in a new representation of Tweets in a space of lower dimensionality. Consequently, the set of Tweets, originally represented through the occurrences of words in their texts, are now represented according to how much they discuss each of the latent semantic concepts, adopting a representation that is no longer based on individual words but rather on mixes of semantically similar terms. Once the LSA space has been selected as the most suitable embedding space in the application, the second step of the analysis can be performed. Indeed, embedding Tweets in the LSA space provides the chance to describe every single text in terms of the concepts it discusses. However, we would like to analyse the entire set of Tweets together, to understand if there are any similarities between the visitors' impressions. To do so, we perform a cluster analysis, which, given a certain similarity measure, aims at grouping similar Tweets together. As underlined before, the novelty of the approach lies not in the well-known cluster analysis technique or in useful LSA, but in reframing the statistical text analysis in the object-oriented data analysis perspective. Thus, it offers a conscious statistical view of the text datum.

4.3 Results

In order to explore the perceptions of visitors to the Queen Elizabeth Olympic Park, this paper shows the application of latent semantic analysis to the set of 279 Tweets written in the English language and published between 16 March and 3 April 2018 from Twitter users located in two specific regions of the QEOP (Figure 3).

Before applying the semantic representation technique to these Tweets, a preliminary step of text pre-processing was applied to define a vocabulary of semantically meaningful terms based on the text of the considered messages. This phase consisted of different authomatic steps, among which the division of the text into individual words (segmentation), the exclusion of terms that were not characteristics for the application setting (stopwords removal), and the reduction of words to their grammatical roots (stemming) led to the construction of a vocabulary of 139 unique words extracted from the Tweets' text. This implies that, according to the vector space model, each of the 279 Tweets was represented as a vector of dimension 139, where most of the entries were zeros.

The application of LSA to the Tweets resulted in the identification of four latent semantic concepts. This allowed for a new representation of the Tweets in a four-dimensional subspace derived from the vector space generated by the 139 unique words in the vocabulary. Indeed, the four main directions of this latent semantic space were particular linear combinations (i. e., a mix of words) of the original 139 terms, whose interpretation allowed individuation of semantic topics of discussion in the Tweets.

Out of the 139 unique terms constituting each of the four mixes of terms, the most influential words in each combination of words helped with interpretation of the corresponding latent semantic concept. In particular, we called the first semantic concept "London Stadium", since the characterizing words were related to events taking place at the stadium. The second mix of words instead mostly combined words associated with the stadium infrastructure and terms related to the Queen Elizabeth Olympic Park, leading to the name "London Stadium & QEOP" for the second latent semantic concept. The third topic of discussion among considered Tweets instead showed a contrast between the park and the stadium infrastructure, highlighting a difference in the meaning expressed by some words related to the QEOP with respect to the adoption of words related to the park infrastructure. Due to this contrasting behaviour, we addressed the third latent semantic concept as "QEOP vs. London Stadium". A similar reasoning held for the fourth concept, called "Westfield Mall vs. London Stadium", which individuated words strongly connected to activities taking place in the Westfield Stratford City Mall in opposition to terms highly related to stadium events.

Following the previously introduced nomenclature for the four latent semantic concepts, the meaning of each Tweet could be expressed in terms of how much the text of the message was related to each of these four arguments.

For example, consider the Tweet: "*Great afternoon at the London Stadium. A big win for West Ham in their bid to avoid relegation this season. #WHUSOU #COYI*". This message is clearly concerned only with a football match at London Stadium and this is reflected in the components of its vector representation. Indeed, the Tweet achieves high weights along the first two concepts of "London Stadium" and "London Stadium & QEOP" and shows high and negative scores in terms of both the third and fourth latent semantic concepts, respectively "QEOP vs. London Stadium" and "Westfield Mall vs. London Stadium", highlighting that the considered Tweet only uses words semantically related to the stadium infrastructure.

Both the data filtering and the latent semantic analysis procedure highlight the importance of a bias-conscious approach to data. On the one hand, the data filtering shows the importance of dealing with a potentially smaller but more accurate sample. Even if this consistent reduction of the sample of Tweets may seem a huge loss of available information, the authors encourage the reader to meditate on the knowledge provided by this subset of Tweets. Indeed, statistically speaking, even if the sample is small, it represents contents published by a specific subset of the population of visitors to the Queen Elizabeth Olympic Park, in particular English-speaking Twitter users sharing their geographic position. On the contrary, the originally gathered Tweets included opinions of a wider population of Twitter users, who agreed on sharing their location but did not necessarily visit the park. Even if strong, this assumption aims to extract knowledge on the specific application framework starting from the available information collected, giving a central role to the application need, considered by the authors the key point of the entire analytic procedure. Therefore, considering that the leading ingredient for the analytic process here adopted was the analysis of *perceptions of visitors to the QEOP*, we decided to discard a sample of Tweets with a large sample size, which contained a lot of information but not necessarily related to the application, in favour of a sample of Tweets with a small sample size, but richer in knowledge with regard to analysing perceptions of visitors to the QEOP.

On the other hand, thanks to the LSA procedure, the data were embedded in a smaller but meaningful space. This latent semantic space gave an easy chance of interpreting the embedded data and the following results, while the interpretation would have been harder in the bigger vector space associated with the Bag-of-Words model.

Once the texts were embedded in this four-dimensional semantic space, we performed a cluster analysis of the text embedding of Tweets, employing a non-hierarchical method (k-medoids) for the grouping of semantically similar messages based on the cosine distance (RIVA 2018). Looking only at the text of messages, hence to the scores of each Tweet along the four latent semantic concepts indivuated through LSA, we found the existence of four clusters of Tweets, i. e., four different impressions, each of which characterized a specific subset of messages. The analyses showed that, during the specific time interval between 16 March and 3 April 2018, there were two different clusters of QEOP's visitors discussing on Twitter events taking place at the London Stadium. In particular, one group of Tweets talked about a Premier League football game[4] while the other discussed over an important Aviva Premiership Derby rugby fixture[5]. Another cluster of Twitter users visiting the park instead mostly published contents related to shopping and strongly connected to the Westfield Mall. Finally, the analyses also individuated a group of visitors who shared on Twitter impressions directly related to the Queen Elizabeth Olympic Park. Observe that, even if the groups of Tweets had not been created by dividing contents according to their geographic location or to the publishing time, the results of the analyses show a clear correspondence between the group a Tweet belongs to and its location (Figure 5).

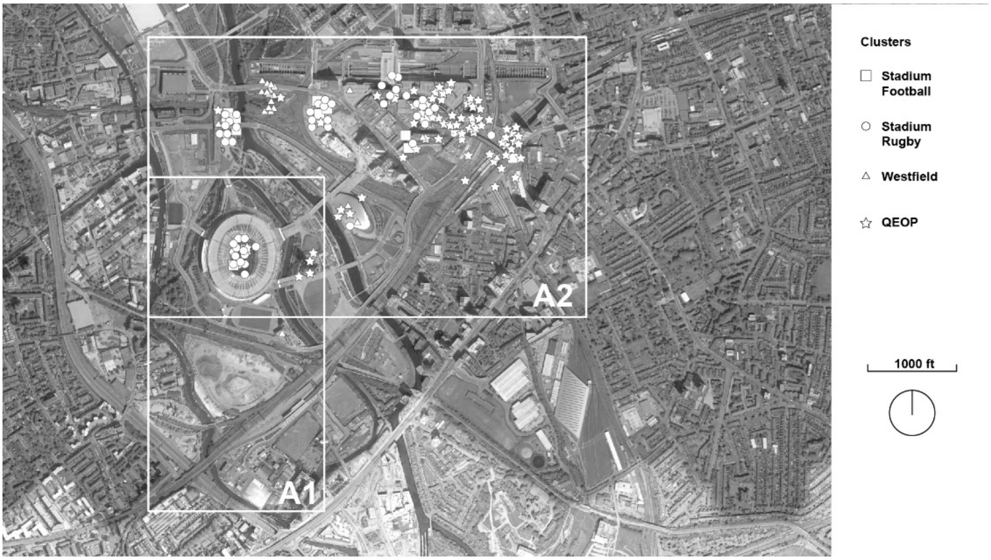

Fig. 5: Location of Tweets inside the two regions of interest in the QEOP on Google Earth aerial image. Symbols corresponds to clusters of semantically similar Tweets

5 Discussion

These results show the existence among considered Tweets of four main semantic areas of discussion, which connect in different ways the park to two main infrastructures towering the region of interest for the analysis, namely the London Stadium and the Westfiled shopping center. This result suggests that the users' perceptions of the Queen Elizabeth Olympic Park consider at the same level the park, the Stadium, and the shopping center. This information is interesting if we consider that Westfield is the biggest mall in the Greater London area and that the Olympic Stadium hosts international events. In a standard transit-oriented development (TOD) framework, the mall and the stadium would be considered attractors and singu-

[4] https://www.whufc.com/news/articles/2018/march/30-march/west-ham-united-v-southampton-all-you-need-know

[5] https://www.queenelizabetholympicpark.co.uk/whats-on/events/2018/03/aviva-premiership-derby-day-fixture-saracens-v-harlequins

larly conceived as efficient machines for conveying flows, the endpoints of pedestrian desire lines originating in the nearest intermodal node. For this reason, the highest budget is commonly allocated to design such compounds, leaving the surrounding landscape in the background. However, the results of this study show that, from the user's perspective, a well-designed, integrated master plan that includes a landscape intervention could compete with the traditional attractors in contributing to shape the experience of the place itself. The results make a case for landscape as a central spine in TODs, to achieve an effective intermodal management of flows. This phenomenon happens because the investment focuses not only on the architectonic quality of the attractors, but also on the area in-between functions, in its materials and forms, making the interface, the connecting spaces, the protagonists of the intervention. That is where the urban life originates, the "vibrancy" (JACOBS 1961), which turns a space into a place.

6 Conclusions

Tweets reveal several differences in the topics discussed following dynamic changes in environmental conditions, rendering the subjective, individual, and evolving impression of the *Umwelt*. As already stated, to impress the user in this regard could be one of the goals of a landscape project, beyond its functional, environmental, and standards-compliancy requirements. These analyses provide a glance into the user's impression and experience of a built space. It is important to stress that this result is not intended as quantitative information, usable for example to evaluate the quality of the design in the context of a due diligence compliance. On the contrary, it constitutes an augmented perception of the place, supporting the designers in sensing the park and observing the recognisability of the designed space.

Moreover, this exploration revealed the potentials of interdisciplinary collaboration as a platform for cross-pollination between different fields of knowledge: it allowed the mathematicians to develop an analytic method based on the research questions posed by the landscape designer. At the same time, the output of the analysis supported the designer in sensing the place. This case of disciplinary cross-pollination goes beyond mere interdisciplinary work, opening opportunities for creative and dense teamwork.

References

ARNABOLDI, M., BRAMBILLA M., CASSOTTANA B., CIUCCARELLI P. &VANTINI S. (2017), Urbanscope: A Lens to Observe Language Mix in Cities. American Behavioral Scientist, 61 (7), 774-793. https://doi.org/10.1177/0002764217717562.

BATESON, G. (1972), Steps to an Ecology of Mind. Jason Aronson Inc., Northvale, NJ and London, UK.

BATTY, M. (2005), Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals. MIT Press, Cambridge, MA.

BATTY, M. (2013), The New Science of Cities. MIT Press, Cambridge, MA.

BOYD, D. & CRAWFORD, K. (2016), Critical Questions for Big Data. Information, Communication & Society, 15 (5), 662-679.

BRUIJN, H. DE & HEUVELHOF, E. TEN. (2008), Management in Networks: On Multi-Actor Decision Making. Routledge, London, UK and New York, NY.

CALISSANO, A., VANTINI, S. & ARNABOLDI, M. (2018), An Elephant in the Room: Twitter Sampling Methodology. MOX report, Politecnico di Milano.

CRANENBURGH, S. VAN & ALWOSHEEL A. S. (2019), An Artificial Neural Network Based Approach to Investigate Travellers' Decision Rules. Transportation Research Part C, (98, January), 152-166. ISSN: 0968-090X. DOI: 10.1016/j.trc.2018.11.014

DEERWESTER, S., DUMAIS, S., FURNAS, G., HARSHMAN, R., LANDAUER, T., LOCHBAUM, K. & STREETER, L. (1988), Computer Information Retrieval Using Latent Semantic Structure. Proceedings of the 11th annual international ACM SIGIR conference on research and development in information retrieval, 01 May 1988, 465-480. DOI: 10.1145/62437.62487.

DEERWESTER, S., DUMAIS, S., FURNAS, G., LANDAUER, T. & HARSHMAN, R. (1990), Indexing by Latent Semantic Analysis. Journal of the American Society for Information Science, 41 (6), 391-407. doi:10.1002/(SICI)1097-4571(199009)41:6<391::AID-ASI1>3.0.CO;2-9.

DEHS, J. (2001), Sense of Landscape: Reflections on a Concept, a Metaphor, a Model. In Olafur Eliasson: Surrounding Surrounded; Essays on Space and Science (Ed. ULRIKE HAVEMANN). MIT Press, Cambridge, MA.

FIRTH, J. (1957), A Synopsis of Linguistic Theory: 1930-1955 in Studies in Linguistic Analysis. Philological Society (Great Britain). Blackwell, Oxford.

HARRIS, Z. (1954), Distributional Structure. Word, 10 (2-3), 146-162.

HOBBS, J., GROSZ, B., HAAS, N., HENDRIX, G., MARTIN, P., MOORE, R., ROBINSON, J. & ROSENSCHEIN, S. (1982), Dialogic: A Core Natural-Language Processing System. COLING 82: Proceedings of the Ninth International Conference on Computational Linguistics, Prague, 5-10 July, 95-100.

JACOBS, J. (1961), The Death and Life of Great American Cities. Random House, New York, NY.

JAKOB, M. (2018), What Is Landscape? Listlab, Barcellona, ES.

JIE, M. & YUNING, C. (2018), Research on the Relationship between Urban Public Space Behavior and Landscape Morphology Based on Big Data of Social Networks. Journal of Digital Landscape Architecture, 3-2018, 356-364. doi:10.14627/537642038.

KANT, I. (1900 [1781]), Critique of Pure Reason. Colonial Press, New York, NY.

KANT, I. (1992 [1791]), Critica Del Giudizio, 6th edition. Laterza, Roma and Bari, IT.

KAUSSEN, L. (2018), Landscape Perception and Construction in Social Media: An Analysis of User-generated Content. Journal of Digital Landscape Architecture, 3-2018, 373-379. doi:10.14627/537642040.

KITCHIN, R. (2014a), The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences. Sage Publications Ltd., Singapore.

KITCHIN, R. (2014b), The Real-Time City? Big Data and Smart Urbanism. Geo Journal, 79 (1), 1-14.

KITCHIN, R. & DODGE, M. (2011), Code/Space: Software and Everyday Life. MIT Press, Cambridge, MA.

KWAN, M., DIJST, M. & SCHWANEN, T. (2007), The Interaction between ICT and Human Activity-Travel Behavior. Transportation Research Part A: Policy and Practice, 41 (2), 121-124. https://doi.org/10.1016/j.tra.2006.02.002.

M'CLOSKEY, K. & VANDERSYS, K. (2017), Dynamic Patterns. Visualizing Landscapes in a Digital Age. Routledge, London, UK.

MANFREDINI, F., PUCCI P. & TAGLIOLATO, P. (2016), Mobile Phone Data in Reading Mobility Practices. In Understanding Mobilities for Designing Contemporary Cities (Eds. PAOLA

PUCCI and MATTEO COLLEONI) Springer International Publishing, Cham, DE, 253-272. https://doi.org/10.1007/978-3-319-22578-4_14.

MANYIKA, J., CHUI, M., BROWN, B., BUGHIN, J., DOBBS, R., ROXBURGH, C. & HUNG BYERS, A. (2011), Big Data: The Next Frontier for Innovation, Competition, and Productivity. Report. McKinsey Global Institute. https://www.mckinsey.com/~/media/McKinsey/Business Functions/McKinsey Digital/Our Insights/Big data The next frontier for innovation/MGI_big_data_exec_summary.ashx.

MARRON, J. S. & ALONSO, A. (2014), Overview of Object Oriented Data Analysis. Biometrical Journal, 56 (5), 732-753. doi:10.1002/bimj.201300072.

MONTANO, F. (2018), The Use of Geo-Located Photos as a Source to Assess the Landscape Perception of Locals and Tourists; Case Studies: Two Public Open Spaces in Munich, Germany. Journal of Digital Landscape Architecture (3), 346-55. doi:10.14627/537642037.

NIKAYIN, F. & DE REUVER, M. (2012), Governance of Smart Living Service Platforms: State-of-the-Art and the Need for Collective Action. Third International Engineering Systems Symposium CESUN, June, 18-20. http://cesun2012.tudelft.nl/images/1/12/Nikayin.pdf.

PORTUGALI, J. & HAKEN, H. (2000), Self-Organization and the City. Springer, Berlin Heidelberg, DE.

PORTUGALI, J. & HAKEN, H. (2011), Complexity, Cognition and the City. Springer, Berlin Heidelberg, DE.

PORTUGALI, J., MEYER, H., STOLK, E. & EKIM, T. (2012), Complexity Theories of Cities Have Come of Age: An Overview with Implications to Urban Planning and Design. Springer, Berlin Heidelberg, DE.

PUCCI, P. (2016), Mobility Practices as a Knowledge and Design Tool for Urban Policy. In Understanding Mobilities for Designing Contemporary Cities (Eds. PAOLA PUCCI and MATTEO COLLEONI). Springer International Publishing, Cham, DE, 3-21. https://doi.org/10.1007/978-3-319-22578-4_1.

RATTI, C., PULSELLI R. M., WILLIAMS, S. & FRENCHMAN, D. (2006), Mobile Landscapes: Using Location Data from Cell Phones for Urban Analysis. Environment and Planning B: Planning and Design, 33 (5), 727-748. http://journals.sagepub.com/doi/abs/10.1068/b32047.

RIVA, P. (2018), Text is a Complex Datum: An Object Oriented Data Analysis of Tweets in Queen Elizabeth Olympic Park. Thesis for Master of Science in Mathematical Engineering, Politecnico di Milano. https://www.politesi.polimi.it/handle/10589/141731.

ROSNAY, J. DE (1979), The Macroscope: A New World Scientific System. Harper & Row, New York, NY.

ROTH, M., HILDEBRANDT, S., RÖHNER, S., TILK, C., SCHWARZ VON RAUMER, H., ROSER, F. & BORSDORFF, M. (2018), Landscape as an Area as Perceived by People: Empirically-based Nationwide Modelling of Scenic Landscape Quality in Germany. Journal of Digital Landscape Architecture, (3), 129-37. doi:10.14627/537642014.

SALTON, G., WONG, A. & YANG, C. (1975a), A Vector Space Model for Automatic Indexing. Communications of the ACM, 18 (11), 613-620. doi:10.1145/361219.361220.

SALTON, G., YANG, C. & YU, C. (1975b), A Theory of Term Importance in Automatic Text Analysis. Journal of the American Society for Information Science, 26 (1), 33-44. doi:10.1002/asi.4630260106.

SAMUEL, A. L. (1959), Some Studies in Machine Learning Using the Game of Checkers. IBM Journal, (3), 210-229. doi: http://dx.doi.org/10.1147/rd.33.0210.

SIMON, H. (1969), The Science of the Artificial. MIT Press, Cambridge, MA.

STILGOE, J. R. (2015), What Is Landscape? MIT Press, Cambridge, MA.

TULLOCH, D. & IM, W. (2018), Towards Using Social Media as a Geospatial Tool for Measuring Design Impact on Human Experiences. Journal of Digital Landscape Architecture, 3-2018, 227-233. doi:10.14627/537642024.

TWITTER INC. (2018), Developer Documentation. https://developer.twitter.com/en/docs.

TWITTER INC. (2018), Geo Objects. https://developer.twitter.com/en/docs/tweets/data-dictionary/overview/geo-objects.html.

TWITTER INC. (2018), Twitter Streaming API. https://developer.twitter.com/en/docs/tutorials/filtering-tweets-by-location.

UEXKÜLL, J. (1926), Theoretical Biology. Kegan Paul, Trench, Trubner & Co. Ltd., London, UK.

VANTINI, S. (2018), Wishing the Non-parametric Re-evolution. Statistics & Probability Letters, 136, 139-141.

WANG, H. & MARRON, J. (2007), Object Oriented Data Analysis: Sets of Trees. Annals of Statistics, 35 (5), 1849-1873. doi:10.1214/009053607000000217.