Transforming our World through Universal Design for Human Development
I. Garofolo et al. (Eds.)

doi:10.3233/SHTI220822

The Missing Voices of Disabled People

Masashi KAJITA[a1] and Emil BALLEGAARD[a]
[a] *Royal Danish Academy, Institute of Architecture and Design*

Abstract. This paper presents the development process of a graph database that connects statements posted by disabled people on various web-based platforms with accessibility requirements in the Danish Building Regulations (BR18). The aim is to bring the lived experience of disabled people into a vocabulary of space-making for architects. By elevating the missing voices of disabled people – describing what matters, how and why – the project supports the decision-making processes of architects to make the built environment more inclusive. The developed database relates statements posted by disabled people with sentences from paragraphs of BR18 through specific architectural features of room, element, and object. Using the architectural features as point of reference, the database not only highlights some of the most common building situations encountered by disabled people, but also allows anyone interested to explore their relationship to the real lives of disabled users and the statutory requirements.

Keywords. Accessibility, Disability, Danish Building Regulations, User Statements, Graph Database

1. Introduction

Architecture still excludes disabled users by not complying with their basic needs and preferences [1]. This is despite more than 1 billion people worldwide living with some form of disability [2]. In recent decades, increasing emphasis on providing equal opportunities for users has improved the inclusion of disabled users in the built environment, but there is still a long way to go [1]. Codified rules and building standards have played a significant role in the reduction of inequalities. However, these rules and standards abstract the relations between humans and the built environment and simplify the complexity of use, leading design solutions to be more generic and standardized [3, 4, 5]. In response to this, several design concepts, such as Universal Design, Inclusive Design and Design for All, have emerged. They all advocate for inclusion through qualitative understanding of human-environment relations. However, their applications to design are still limited in present architectural workflows [6].

Broadly speaking, there are two different approaches to tackling the urgent problem of access disparity in design, both of which have drawbacks. The first is rule-based, drawing on the objectified human body and specified standards, the other is user-based, being characterised by embodied and highly subjective understandings of architectural experience. What we strive for is a strategy that combines the effectiveness and impact of regulatory frameworks with the qualitative approaches found in the people-centred

[1] Corresponding author; masashi.kajita@kglakademi.dk

design concepts [5, 7]. In the effort of providing practical foundation for such a strategy, this paper presents the initial development process of a graph database that connects statements posted by disabled people on various web-based platforms with accessibility requirements in the Danish Building Regulations (BR18). Employing mixed-methods that accurately draw on quantitative scalability and qualitative descriptions to adequately transpose user experiences, the project seeks to provide architects with insights into specific aspects of users' experience in the build environment. Increased access to data and the proliferation of database technology constitute an opportunity to support the decision-making process of architects to make the built environment more inclusive.

2. Background

Codified rules and standards have played a significant role in reducing inequality of access and usage in the built environment by defining minimum obligations and basic architectural performance measures. The Danish building regulation BR18 [8] prescribes requirements and guidelines in various parts of the built environment. The requirements for residential buildings are described in BR18 as most common building situations encountered by users in their everyday lives. Situations such as 'access' (§48-62), 'kitchen work' (§202) and 'outlook' (§378) are articulated through specific requirements that are formulated as descriptions and measurements of architectural elements, spatial dimensions, and/or minimum performances required. The codified rules and standards, based on generalized and decontextualized expert knowledge, provide the essential point of reference for architects. And yet, they suppress the differences that human bodies possess and simplify the complexity of bodily interactions with objects, buildings, and their surrounding environments. These shortcomings raise questions about how well these rules and standards can respond to the real experiences of users [4, 5].

As opposed to this top-down approach of regulations, several design practices have emerged, which advocate for inclusion through qualitative understanding of human-environment relations. Though these approaches vary in name - such as Universal Design (UD), Inclusive Design or Design for All - they all promote designed environments that are more responsive to diverse bodily differences and situations [9]. These practices engage with a broader group of users, through which they contribute to making the built environment more accessible and inclusive [6, 10]. However, their instrumentality on design is often questioned - especially when they are codified into legislation [3, 5, 11]. Debate continues over the applicability of people-centered methods, as they focus on highly personal and subjective bodies, which could lead design solutions to be too specific to the selected individuals. To gain a holistic understanding of human behavior and to develop evidence-based knowledge that is applicable in architectural praxis, we must systematically combine findings from different studies.

In recent times, the availability of user data and a growing awareness of user-environment relationships has created renewed interest in data-based user perspectives in architectural design [12, 13]. Parallel prevalence of tools to collect, process, and store data has made data-based descriptions of user experience accessible for many [14]. This we take as an opportunity to combine the qualitative approach of people-centred methods based on data from disabled people with the quantitative approach of legislation. In what follows, the paper describes how we developed a graph database that relates the situated

experience of disabled users – statements posted on various web-based platforms – with accessibility requirements in BR18.

3. Methods

3.1. Data collection

The data deployed in this project was scraped from personal blogs written by disabled people. Our criteria for the selection of blogs were: 1) the blog is written in English; 2) the blog is written by wheelchair users whose content primarily consists of personal matters; 3) the blogposts are open to the public and the blogpost is not considered sensitive or exposing. We chose English due to highly limited numbers of Danish bloggers. We focused on wheelchair users due to the large concentration of blogs made by wheelchair users in virtual space. This focus aligns with the fact that the guidance regarding accessibility in building regulations is largely based on the capabilities of independent wheelchair users [15]. We discovered 108 blogs from all over the world that cover a variety of topics: some post about food, fashion or travels – while others post about activism, culture, or use the blog as a public diary. Even though the topics are divergent there is a common focus – sometimes explicit, sometimes very subtle – on the 'life as a wheelchair user' with personal stories and a high degree of ecological validity.

Although many of the blogs are built as webpages using standard templates from web services like WordPress, there are always unique aspects which challenge the application of a one-size-fits-all scraping solution. For this reason, and for the sake of easy replicability, we tested three 'off-the-shelf' browser-extension scraping solutions. Although we had the best results with Web Scraper, an extension for Google Chrome, its performance was still limited; it successfully scraped just 57 of the 108 blogs. The main problem in the scraping process was entries based on multiple links, dynamically loaded pages and inconsistent URLs and pagination links. The outputs from the scraping process are the blogpost URL, the author of the blogpost, and the actual blogpost content. 4165 blogposts (with text equivalent to 6200 standard pages) were scraped from the 57 blogs. The extent of data from the different blogs varies: some bloggers wrote more than 500 blogposts, while others wrote fewer than 5 blogposts.

3.2. Natural Language Processing (NLP)

The raw data is first cleaned and fragmented into individual tokens and metadata is subsequently extracted for each token. The transformation and creation of metadata involves a number of sub-processes. These processes are carried out using an open-source Python Natural Language Processing (NLP) library called spaCy [16]. The scraped data is highly disordered, with HTML code, image links and other irrelevant information. Therefore HTML and ASCII character filters are deployed in order to filter out irrelevant and infrequent characters from the scraped data. The blogposts are then divided into individual sentences using the spaCy syntax-driven sentence segmentation feature. Following this is a process of tokenization, which breaks the text/sentences down to single tokens. Words that appear very frequently in the English language, but which do not carry any relevant information, such as "the", "at", "is", "on" and so forth, are called stop words. The standard list of stop words provided in spaCy is used to remove

them, which helps to minimize the extent of irrelevant data in future searches, as a search for a word like "the" would otherwise result in many hits with highly varied content.

In linguistics, the concept of dependency refers to the syntactical structure of directed links between words [17]. The keystone of clause structure is the finite verb; all other words have either direct or indirect links to it, which are known as dependencies. For example, in the sentence "The door was wide and accessible", the word "accessible" is directly related to the word "wide" and serves as a conjunct. The parsed dependencies reveal the syntactic relation among the tokens and provide an understanding of how the same word can be used differently in various contexts. Dependencies also reveal who or what is the grammatical subject and object of a sentence. The subject is often found as the individual or thing that is performing the action and is often related to the main verb of the sentence. Objects are often found as the receivers of the action in a sentence. Knowing the subject and the object can provide information to distinguish relevant sentences from irrelevant.

In natural language, anthimeria (the usage of a word in a different grammatical forms) is relatively common and can cause confusion, both in terms of understanding the meaning of the word, and as the cause of incorrect statistics. For example, the word "book" is used both as a noun "a book" and as a verb "to book a ticket". Part-Of-Speech-tagging (POS-tagging) identifies the corresponding part-of-speech of each word. The entire dataset that contains statements from blogposts and paragraph sentences of BR18 is processed, and each word is stored with its dependency relation and POS-tag as metadata. The metadata is used in the following process described in 3.3, both to define the relations between sentences and words, and to create the index categories based on architectural features. By developing such a structure, we seek to enhance the usability of the database for non-specialists with limited previous understanding of linguistic labels.

3.3. Graph Database

A Graph Database (GDB) is a database that uses graph structures consisting of nodes and edges for storage of data and for semantic queries [18]. A node is comparable to a spreadsheet cell and represents an entity or an instance. Nodes are labelled according to predefined sets and can hold different datatypes that describe the entity it represents as key-value properties. Edges are labelled relations between the nodes and can also have assigned properties. A central aspect of the GDB schema design is the distinction between data stored as nodes, edges and properties. The GDB is chosen due to the main objectives of this project, which are to: 1) handle complex and interconnected data which challenge table-based relational databases [19], 2) create meaningful relationships between different data sources and 3) present the interconnected data in a way that it is usable for third parties through web-based publishing. The processed blog data and the paragraphs of BR18 are loaded into a NEO4J GDB according to the schema in Fig. 1.

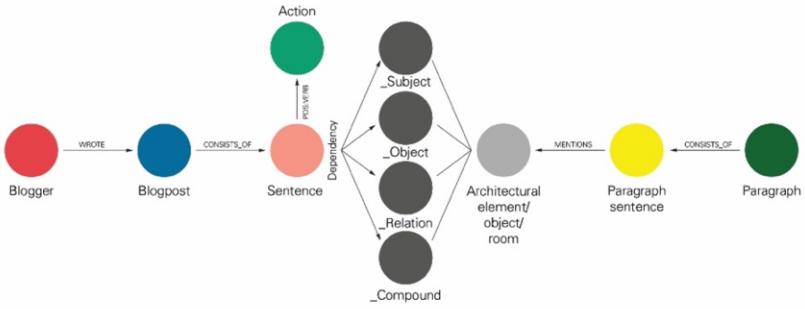

Figure 1. the Blogger-node

Starting from the left in Figure 1: the Blogger-node holds the name used by the blogger. The Blogpost-node contains the URL of the actual blogpost. The blogpost is stored in the database as individual Sentence-nodes that contain each individual sentence in the blogpost as key-value properties. The Action-node contains verbs used in the sentence that match a list of predefined verbs found potentially relevant to architectural context and use. Examples of these "actions" would be verbs such as, "move", "sit" and "see", among many others. The four dark grey coloured nodes represent how features in the Architectural element/object/room-node are used syntactically. If an architectural element such as "wall" is mentioned in a sentence, the connection between that specific sentence-node and the architectural element-node named "wall" describes its use. Besides "subject" and "object" described in the above sub-section 3.2, two additional categories are made for the syntactic use of words as "relation" and "compound".

A relational use of "wall" would be in a sentence that describes the wall in relation to something else, e.g.: "…the space between the *table* and the *wall* was too narrow.". A connection through a compound-node shows that the word is used in combination with others, such as "a *wall mounted* handrail". Architectural element/object/room consists of 58 pre-selected elements such as wall and floor, objects such as door and chair, and functionally defined rooms such as bedroom and kitchen. These nodes provide an overview and allow users of the database to compare how the architectural features are mentioned across the sentences from both blogposts and BR18 paragraphs. The asymmetric relationship between blogpost-sentences and paragraph-sentences of BR18 emphasizes the differences between the legislative information and narrative information describing embodied experience. Much like their Sentence-node counterparts, Paragraph-sentence-nodes contain individual sentences from the entire paragraphs of BR18. These paragraphs are, in turn, represented by Paragraph-nodes that hold a URL to that specific paragraph published on the official BR18 website.

3.4. Querying/using the database

The graph database is published online [20] and can be queried using three different search methods as seen in Fig. 2. The first of these is by using the search bar, where literal string searches can be conducted. This means that the search engine only matches the search results when the spelling and the word order are the same. The second is by choosing the nodes of interest from a list containing the different node types described in 3.3. Different nodes can be combined to specify the search. This feature can be used to create abstract narratives, using the linguistic labels in combinations with the

architectural features. As an example, if an architect is designing an entrance hall and wants to know if any blogger has described specific situations related to an entrance, a search could be combined of: (1) Entrance (subject), 2) Door (subject), 3) Floor (object), 4) Wall (object), 5) Wall (relation), 6) Enter, open, and turn (Action/Verb). Sentences that contain any of these words in the specified grammatical form are related to those nodes and it is easy for the user to understand which sentences relate to which nodes, and to see those with the highest number of node relations.

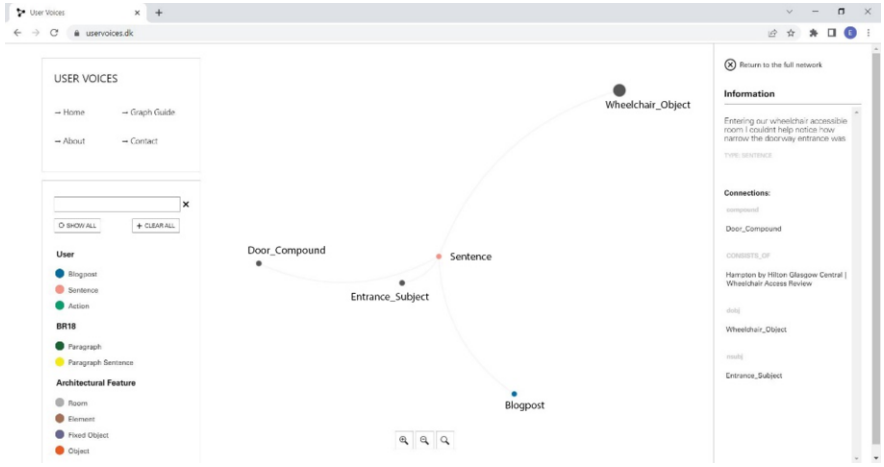

Figure 2. Screenshot from our website uservoices.dk, showing a Sentence-node containing the sentence: "*Entering our wheelchair accessible room I couldn't help notice how narrow the doorway entrance was.*", and its relations to other types of nodes.

The first two search methods can be combined with the third: intuitive exploration of the database by clicking nodes and uncovering their relations. This approach can be used to expand the initial search in unplanned directions guided by reading of nodes, relations, and properties. It can be a useful tool to gain a deeper understanding of the relations between different nodes through several links, and to reveal when parts of the design task are inadequately described. In addition, the intuitive search approach technique is often used to arrive at the blogpost-nodes and the paragraph-nodes, where a link to the actual blogpost or the specific paragraph at the BR18 website can be found. Here, both sources of information can be further investigated in their original context.

4. Results

In this project, we creatively explore a new way to develop knowledge of inclusion that enables architects to gain more nuanced understanding of spatial experiences. The project approaches this challenge by visualising data sets collected from web-based platforms that are highly personal (qualitative) and by establishing relations to the regulatory definitions of BR18 (mainly quantitative). The developed database – presented as a website - is an alpha-version of a framework based on Web Scraping, Natural Language Processing and a Graph Database System that integrates profound subjective perspectives on the everyday lives of disabled people into a vocabulary of space-making for architects. The website is supplemental to BR18 and should support

architects in making appropriate decisions in their planning and design processes of inclusive architecture. Accordingly, it adds to, and must be read in conjunction with, the existing body of regulations that ensure the built environment functions in the interests of public health and safety – by defining basic performance standards for design, construction, and alternation.

The database stores paragraphs of BR18 and statements from disabled users which can be accessed through our website. The structure of the graph database and the architectural index categories provide an overview of which architectural features are mentioned, and where, across the dataset. This makes it easy to find the paragraphs and/or user statements that include specific combinations of architectural features and thereby to identify relevant personal and regulatory information for specific design processes. The database structure connects the paragraphs with the user sentences through shared architectural features. These connections support investigations of the personal and regulatory aspects of common building situations encountered by disabled users. It is important to state that the indirectness of the connections underlines that paragraphs and user-sentences that are related through the same words by no means necessarily have anything in common on a semantic level besides the shared word. However, we, who have professional architectural understanding, concluded that a majority of the linked statements add a potentially relevant personal perspective to the paragraph sentences of BR18 – and that additional useful information often can be found in the blogposts where the sentences originate.

Not surprisingly, we found that the database contains a far greater diversity of situations described by the bloggers than those covered in BR18. This emphasizes that the database can be used to gain insight into specific real-life situations experienced by users that are not covered by regulatory definitions. Exploring the database for blogpost content with few or no node relations to BR18 paragraphs could help to identify these missing regulatory aspects. The relations found in the database might also reveal discrepancies between user stories and regulations relating to the same architectural feature; as such, it could be a tool for the evaluation and improvement of regulations.

5. Conclusion

The project is at an early stage of development and while results seem promising, we are not in a position to draw any justified conclusions on the general validity of the data connections and applicability of the system. Further research needs to be conducted in order to finalise the system into a suitable format for architects and other actors in the building industry to utilise effectively. In continuation of this, we must underline that none of the researchers contributing to this project are educated linguists, programmers, or data scientists. We encourage anyone with specialized knowledge within these fields to point out any errors, suggest changes or reach out for potential cooperation on similar projects.

The database system presented in this paper can be used to query for specific experiences and situations encountered by wheelchair users with or without relation to paragraphs of BR18. The queries can be conducted to gain insights into the narratives of the wheelchair users and to understand what matters to them, how, and why. Moreover, queries can be conducted to find statements that add a personal dimension to the regulatory definitions in BR18 and thereby potentially provide a better understanding of the challenges that the legislation seeks to address. The developed database is a prototype

tool which showcases, at least in principle, that qualitative and quantitative datasets can be structured in such a way as to reveal relations between them. At this point our system is not 'universal' per se and focuses explicitly on a group of wheelchair users, however, it represents a crucial first step towards our aim of combining the effectiveness of regulatory frameworks with the qualitative approaches found in people-centred design concepts into a coherent design strategy.

Acknowledgements

The authors gratefully acknowledge the support of the Bevica Foundation and would like to thank all colleagues involved in the platform "Visualizing Sustainability" that aims to demonstrate new approaches to, and research in, data visualization in design and architecture at the Royal Danish Academy.

References

[1] UN. Accessibility and Development: environmental accessibility and its implications for inclusive, sustainable, and equitable development for all. (2013)
[2] WHO. World Report on Disability. (2011)
[3] Grangaard S, Frandsen AK. "Do Performance-Based Codes Support Universal Design in Architecture?" in: Studies in Health Technologies and Informantics. 2016;229:98-100
[4] Imrie R, Street E. "Architectural Design and Regulation", Wiley, London (2011)
[5] Kajita M, "What Kind of Bodies Do Architects Imagine When Designing?", in: Hauderowicz D, Serena, KL. (editors) *Age-Inclusive Public Space*, Hatje Cantz Verlag, Berlin (2020), 116-119
[6] Bringolf J, "Calling a Spade a Shovel: Universal, accessible, adaptable, disabled – aren't they all the same?" in 2009 Housing Researchers Conference: Refereed conference proceedings. 2009;4
[7] Kajita M. "Inclusion in Danish architectural education and design process.", in: Bucchianico P D (ed.) *Advances in Design for Inclusion,* a chapter in Springer Multi-volume Edited Books under the series of "*Advances in Intelligent Systems and Computing"* (proceedings of the AHFE 2019 International Conference on Design for Inclusion and the AHFE 2019 International Conference on Human Factors for Apparel and Textile Engineering, Washington D.C., USA, July 24-28, 2019). 3-14.
[8] www.bygningsreglementet.dk
[9] Clarkson J, Coleman R, Keates S, & Lebbon C (Eds.) "*Inclusive design: Design for the whole population.",* Springer, London (2003)
[10] Sørensen R, Ryhl C, "Responding to Diversity Including Disability" in: Proceedings of Design Research Society 2018 International Conference: Catalyst. 2018;5:1894-1908
[11] Bianco L, "Universal design: From design philosophy to applied science", in: Journal of Accessibility and Design for All. 2020;10:70-97
[12] Dalton R C, Derix C, "Woods Bagot: user-experience typologies of space", in: Hensel M U, Nilsson F (eds.) "The Changing Shape of Practice" 2016;36-46
[13] GXN (Firm) "Transparency Toolbox" (2020)
[14] Dalton C, Wilmott C, Fraser E, Thatcher J. ""Smart" Discourses, the Limits of Representation, and New Regimes of Spatial Data", in: Annals of the American Association of Geographers 2020;110:485-496
[15] Goldsmith S, PRP Architects (Firm) "Universal Design: A manual of practical guidance for architects", Architectural Press, Oxford; Boston (2000)
[16] www.spacy.io
[17] De Marnefee, M, MacCartney B, Manning C D, Generating Typed Dependency Parses from Phrase Structure Parses*",* in: Proceedings of the Fifth International Conference on Language Resources and Evaluation. 2006
[18] Robinson I, Webber J, Eifrem E, "Graph Databases", O'reilly Media Inc. (2015)
[19] Brunenberg P, "At Its Core: How is a Graph Database Different from a Relational One?" https://towardsdatascience.com/at-its-core-hows-a-graph-database-different-from-a-relational-8297ca99cb8f (2021)
[20] www.uservoices.dk