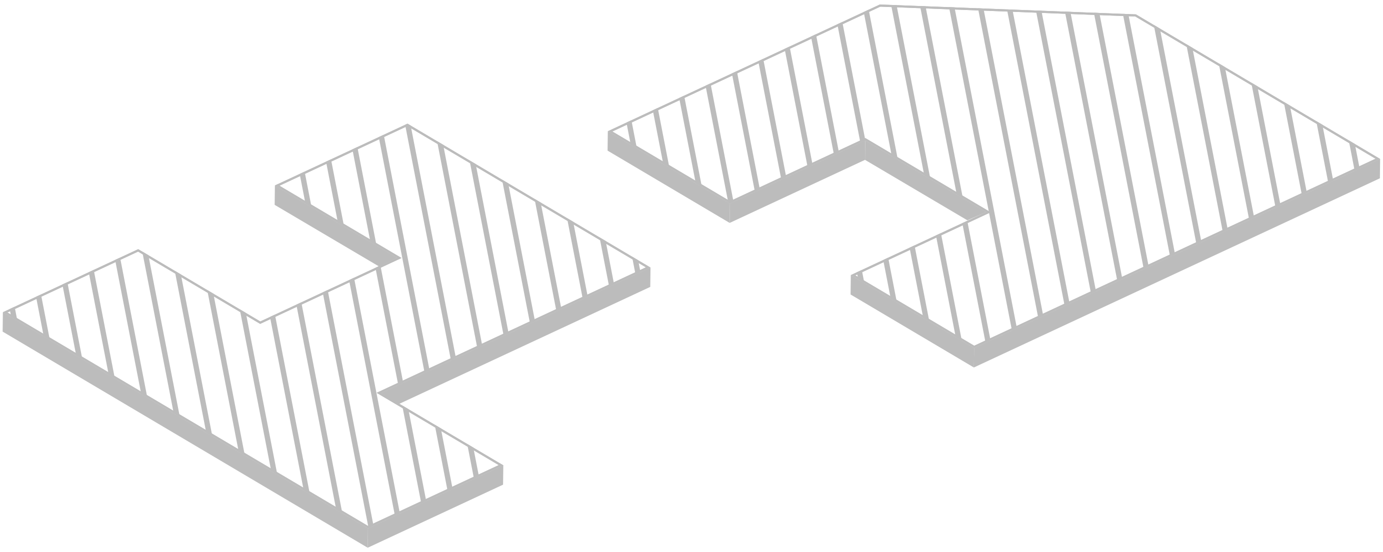
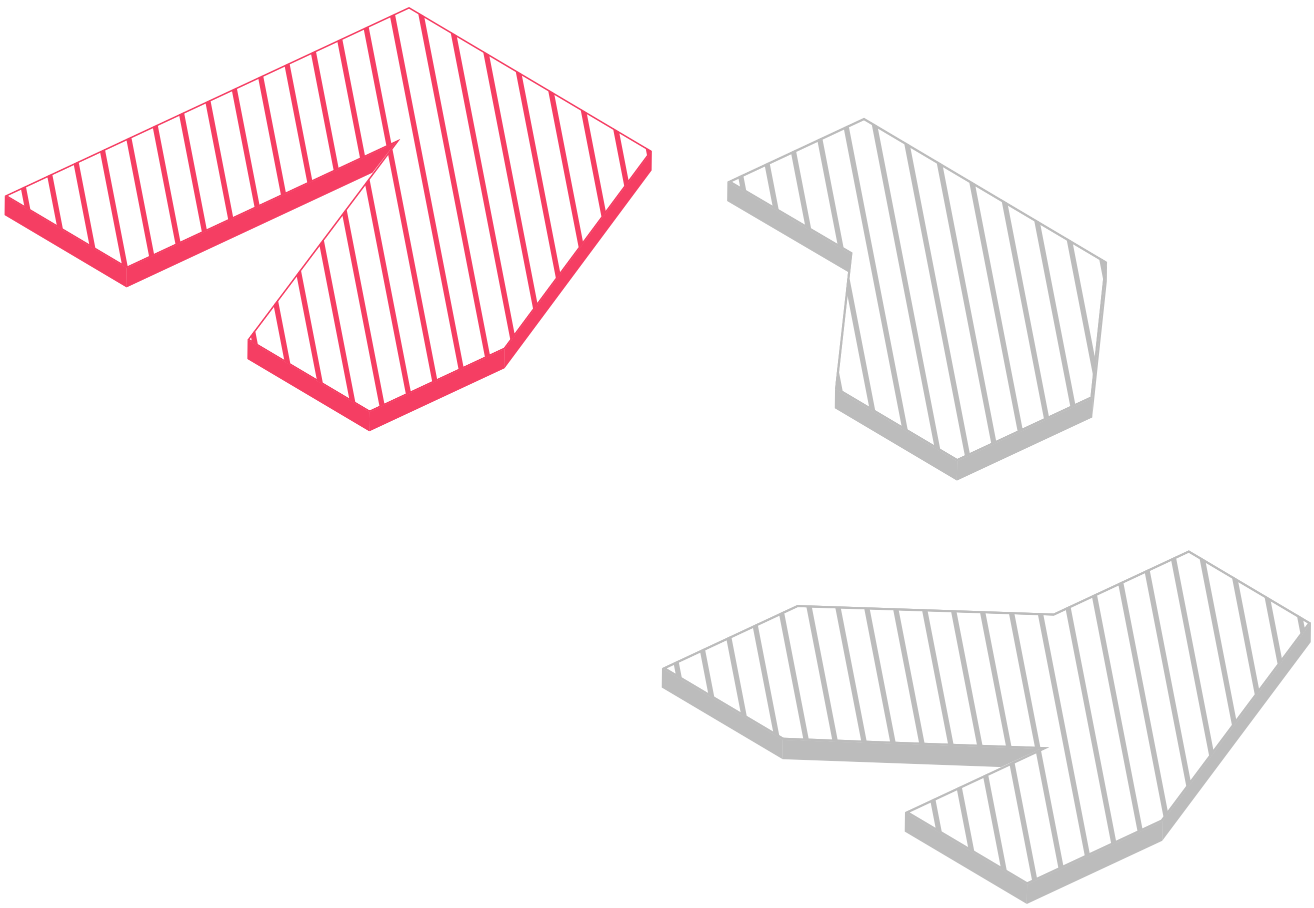


CITY VIS 2018 WORKSHOP



CONTENTS

CREDITS

WORKSHOP CHAIRS

Sebastian Meier, Technologiestiftung Berlin
Sarah Goodwin, Monash University
Nina Hälker, HafenCity University

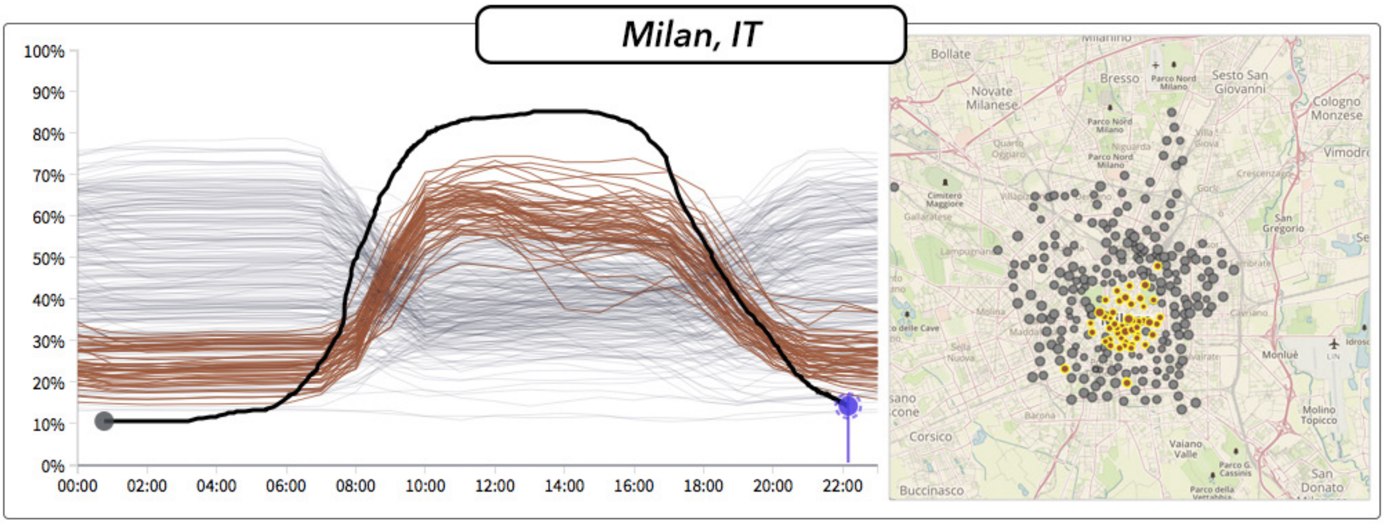
PROGRAM COMMITTEE

Susanne Bleisch, FHNW Northwestern Switzerland
Melanie Davern, RMIT University
Catherine D'Ignazio, Emerson College
Marian Dörk, University of Applied Sciences Potsdam
Manuela Garreton, Pontificia Universidad Católica de Chile
Amy Griffin, RMIT University
Robert S. Laramee, Swansea University
Till Nagel, University of Applied Sciences Mannheim
Astrid Nieße, University Hannover
Jörg Rainer Noennig, TU Dresden
Christopher Pettit, The University of New South Wales
Holger Prang, TU Berlin
Juan Francisco Saldarriaga, Columbia University
Cagatay Turkay, City University of London
Sarah Williams, MIT Media Lab
Jo Wood, City University of London
Wenfei Xu, MIT Media Lab
Gesa Ziemer, HafenCity University Hamburg
Michael Wybrow, Monash University

This endeavor is supported by the Senatsverwaltung für Wirtschaft, Energie und Betriebe and the Investitionsbank Berlin using funds from the Land Berlin.

be Berlin

Investitionsbank Berlin



See Figure 1.

Michael Oppermann[1] (opperman@cs.ubc.ca) & Tamara Munzner[1] (tmm@cs.ubc.ca)
[1] University of British Columbia

UNCOVERING SPATIOTEMPORAL DYNAMICS FROM NON-TRAJECTORY DATA

Keywords: non-trajectory data; spatiotemporal patterns; visual exploration; bike sharing; space utilization

1. DESIGN PROBLEM

In a previous project [5], we analyzed and compared hundreds of bike sharing networks worldwide, predominantly based on station fill levels that we recorded continuously over a period of 17 months. We illustrated how we can support users with a wide range of expertise to understand and intelligently leverage this type of data in their decision-making. Our interactive visualization can reveal interesting insights, not only into patterns of bicycle usage but also into underlying spatiotemporal dynamics of a city (see Figure 1).

We have begun a new project that is focused on creating visual and predictive decision-support tools centered around building occupancy data. Previously this data has been used for the automatic control of heating, ventilation and air conditioning (HVAC) systems and now, by following the design study methodology [6], we are opening it up to a broader set of stakeholders in facility management. Initial experiments indicated that making this data accessible and visually explorable can lead to a better understanding how space is actually being used and will ultimately improve space utilization and resource management.

The intriguing underlying similarity between these projects lies in the data characteristics. We are using status changes at distinct locations (non-trajectory), such as the number of available bikes at a docking station or the number of people occupying a certain room, to investigate spatiotemporal patterns. In contrast, much of the previous work has been focused on individual movements (trajectory) or on origin-destination (OD) data.
By noting the similarity in the data, we can take what we learned in both projects to discuss general implications of spatiotemporal non-trajectory data in terms of ethics, data preprocessing, tasks, and visual encodings. Our goal

is to generalize our findings in the context of urban data visualization with the hope to inspire other researchers and designers.

2. IMPLICATIONS FOR ETHICS

Increasingly digitized cities and massive accumulations of data, often containing geo-tagged personal information, pose risks and raise privacy concerns. Especially fine-grained trajectories can disclose sensitive information, for example, a person's home, workplace and daily commuting pattern. Anonymization and aggregation can only help to a certain degree if the number of trajectories is small. Although there are several methods to address this issue, such as spatiotemporal generalization [3], we should reconsider if the storage and analysis of trajectories is necessary in the first place. We have experienced many barriers and delays when dealing with raw trajectory data and this trend will be reinforced as organizations are implementing the General Data Protection Regulation (GDPR).

On the other hand, generalized and location-based counts are often willingly shared or can be easily accessed in open data libraries.

3. IMPLICATIONS FOR DATA PREPROCESSING

The raw data contains meta information (e.g. geographic location) and a corresponding spatial time series for each sensor. The number of sensors can vary from a few dozens to many thousands. Typically, the sensor states are recorded every few minutes but this can be further shortened to generate more detailed time series.

In our bike sharing project we collected and preprocessed live data from more than 20.000 docking stations in a 15-minute interval. In general, if you have enough knowledge about the data, we suspect that rather than using a uniform interval, you can distinguish between day and night times and adjust the interval accordingly.

4. IMPLICATIONS FOR TASKS

Clearly, the analysis of movement flows is more difficult, if not even impossible, without trajectories or OD data but other tasks are very well supported indeed.

The task of investigating temporal footprints of locations (sensor level) or whole regions (i.e. district aggregation level) is often in the center of the analysis. We can compare sensors at different time resolutions and identify trends, outliers and repeating patterns. Existing approaches are often focused on daily, weekly, and monthly resolutions and lack the ability to capture seasonality. In our latest project, we recognize a strong need from multiple stakeholders for inspecting the data at custom time intervals, for instance, to analyze the utilization of meeting rooms on Mondays between 9am and 5pm. High-level overviews are necessary but they are not sufficient. In addition, users want to see the original time series for a specific location or compare multiple time series on a detailed level. A benefit of this data type is, for the task of capturing local and global variations, that it is much easier to group sensors by regions and to normalize the time series.

Back to the bike sharing example, we realized that the balancing of docking station fill levels poses the biggest challenge for operators. Individual routes taken by customers are usually negligible but operators need to know when a station gets full or empty or what the ideal distance between stations is. They want to understand the distribution of sources and sinks and if behaviors are limited to specific stations or if they are affecting adjacent neighborhoods. We noted many specific questions in this project that could be well addressed without requiring trajectories.

5. IMPLICATIONS FOR VISUAL ENCODINGS

Non-trajectory data can be visually encoded with the same techniques that have been proposed for visualizing time-oriented data. In particular, small multiples, interactive linked views or superimposed perspectives allow us to capture both spatial and temporal dimensions.

In our bike sharing system, we have used an interaction technique to visualize flows indirectly (see Figure 1). Users can draw a custom fill level curve on top of the multi-series line chart and stations that follow a similar pattern are highlighted in the linked map view. By drawing multiple curves and observing the change of active stations, users can get a quick intuition about commuting patterns.

Wood et al. [7] used geographic small multiples to visualize docking stations of London's bike sharing scheme at different points in time. Miranda et al. [2] proposed a multi-view pulse monitor to inspect Flickr activity data. Morphocode [4] discussed different visual encodings for visualizing pedestrian counts in an online blog post.

The obvious common method that people choose when working with trajectories is to display flows as a superimposed layer on a geographic map. This may seem aesthetically pleasing, but the result is often cluttered and difficult to analyze. Although many techniques have been proposed to mitigate this problem [1], we propose an alternative, a data transformation to location/sensor-based counts which would then afford other types of visual encodings.

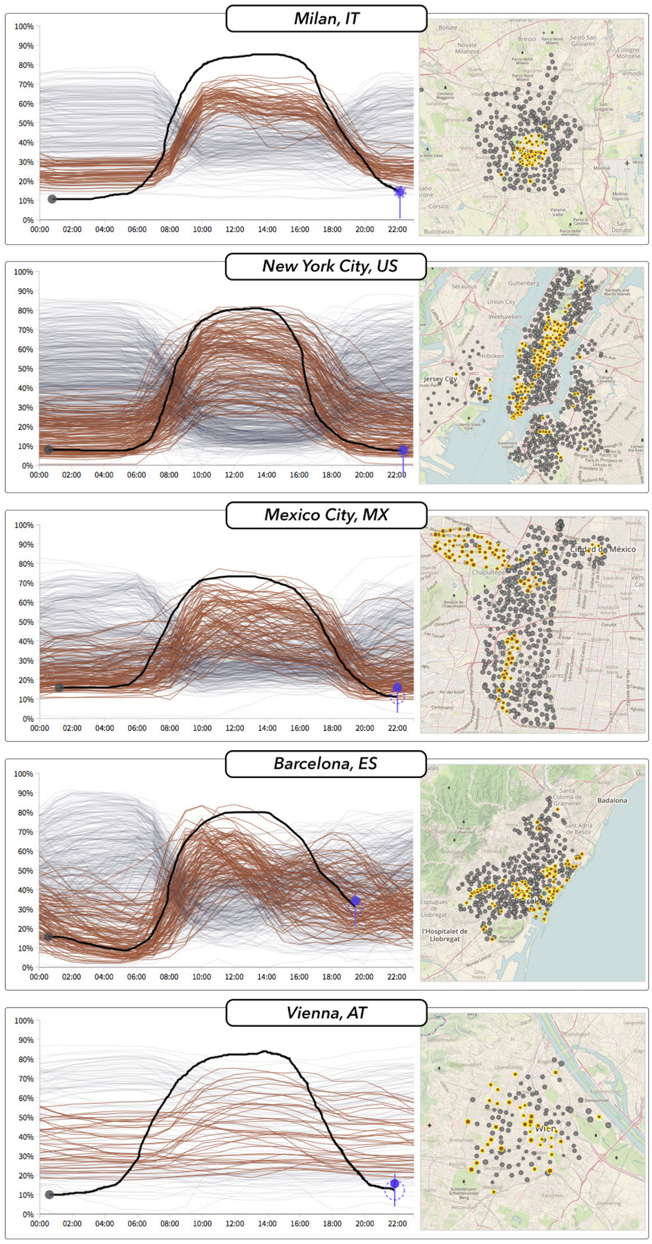


Figure 1: The visualization of average docking station fill levels exposes bike-sharing commuting behaviors in different cities around the world [5]. Each line represents an individual bike sharing station. Users can draw a line—a hypothetical profile—on top of the chart to select similar stations.

REFERENCES

[1] W.Aigner,S.Miksch,H.Schumann,andC.Tominski. Visualizationof time-oriented data. Springer Science & Business Media, 2011.
[2] F.Miranda,H.Doraiswamy,M.Lage,K.Zhao,B.Gon ¨ıalves,L.Wilson, M. Hsieh, and C. T. Silva. Urban Pulse: Capturing the Rhythm of Cities. IEEE Trans. Visualization and Computer Graphics, 23(1):791– 800, 2017.
[3] A. Monreale, G. Andrienko, N. Andrienko, F. Giannotti, D. Pedreschi, S. Rinzivillo, and S. Wrobel. Movement data anonymity through generalization. Trans. Data Privacy, 3(2):91–121, 2010.
[4] Morphocode. Location + time: urban data visualization. https://morphocode.com/location-time-urban-data-visualization/. Accessed: 2018-06-11.
[5] M.Oppermann,T.Mo ̈ller,andM.Sedlmair.BikeSharingAtlas:Visual Analysis of Bike-Sharing Networks and Urban Commuting Patterns Worldwide. Int. Journal of Transportation, 6(1):1–14, 2018.
[6] M. Sedlmair, M. Meyer, and T. Munzner. Design Study Methodolgy: Reflections from the Trenches and the Stacks. IEEE Trans. Visualization and Computer Graphics, 1(12):2431–2440, 2012.
[7] J.Wood,A.Slingsby,andJ.Dykes.Visualizingthedynamicsof London's bicycle-hire scheme. Cartographica: Int. Journal for Geographic Information and Geovisualization, 46:239–251, 2011.

Pascal Goffin[1], Aspen Hopkins[1], Wesley Willett[2], Miriah Meyer[1]
[1]Visualization design lab, University of Utah
[2]Data experience lab, University of Calgary

CHALLENGES IN URBAN AIR QUALITY DATA VISUALIZATION

Keywords: Air quality, sensor network, visualization for the masses

1. INTRODUCTION

Poor air quality (AQ) is one of the most important challenges of increased urbanization. Recent advances in sensing technology have brought low-cost, real-time AQ sensing to cities around the globe [4,9]. Many of these sensors measure fine-particulate matter (PM2.5) — particles with diameters smaller than 2.5 microns — which has the greatest adverse effect on health [1,5,7]. Salt Lake City has a diverse set of PM2.5 sensor networks, including low-cost sensor networks supported by grassroots organizations [6] and government-run networks of high-quality, gold-standard sensors, to measure PM2.5. Currently, the city has no system that unifies all these sensor measurements of diverse sources to provide a comprehensive estimate of PM2.5 levels for its residents [2,8].

2. MEASUREMENT SYSTEM

To meet this challenge, we are developing the AQ&U instrument, a multi-layered tool that collects, unifies and displays PM2.5 measurements from sensors across the city. We collect the data from available sensor networks, while also filling in measurement gaps with our own network of calibrated sensors that are hosted by citizens at their homes. The data from the variety of sensor networks are combined in a modelling layer that generates a continuous estimate of PM2.5 levels across the city, along with measurements of uncertainty. These estimates are visualized in a public-facing web portal to provide real-time information to the public.

3. DESIGN PROCESS

During our formative study, we conducted 6 semi-structured interviews with SLC residents to better understand their motivations to use and assumptions about AQ [3]. These residents were part of a convenience sample but represented a diverse set of people with different socio-economic backgrounds. Based on the results of this study, we built our current prototype to address two use cases: giving sensor hosts access to their AQ measurements and providing us with a testbed to develop our model. When we have gained a better understanding of our model, we will invite multiple sensor holders to participate in a contextual inquiry. We will also include logging to our visualization interface to understand how it is being used.

4. CHALLENGES

In designing the visualization interface, we encountered three challenges that we speculate are general to other environmental sensor endeavours. We are interested in discussing these challenges with other workshop participants.

The first challenge we faced stems from the seasonality of public interest in AQ. Poor AQ occurs mainly in the winter during 1-2 week episodes of meteorological inversions. These weather events trap air in the city and lead to a build-up of PM2.5. During the winter, and particularly during inversions, the public is highly interested in local AQ — but during other times, it is difficult to engage with residents, particularly passively through an instrumented website. This seasonality has non-negligible impacts on us, as researchers, in terms of the timeline to gather data and test the tools, as well as on when to engage the public for ecologically valid formative work.

The second challenge comes from the gulf between what the general public wants to know about AQ and what questions the data can actually answer. When questioning the general public about what they would like to do when gaining access to air quality data during formative work, they responded with a variety of high-level questions. For example, they wanted to understand what they could be doing to improve air quality, such as if buying an electric car would have a beneficial impact on their air quality. Answering these types of questions requires significant analysis of the PM2.5 measurements in combination with simulations, predictions and data from other sources. Simply providing the AQ data is unlikely to meet the public's expectations of what they would like to know, potentially resulting in frustration.

The third challenge involves reconciling the public's mental model for interpreting AQ measurements with the types of measurements our instrument produces. In Salt Lake City, as in many other cities in the US, AQ measurements are communicated based on an EPA classification that associates levels with health impact. For example, green means healthy whereas red means dangerous to health. This classification scheme, however, was not developed for real-time measurements, and thus is not suited for the estimates we produce. It is unclear how to meaningfully, and rigorously, present AQ measurements to the public to support individual decision-making.

5. CONCLUSION

We believe these challenges are likely faced by other visualization design teams working with environmental sensing systems and communication for the public. We are interested in brainstorming with other participants about possible solutions.

REFERENCES

[1] American Lung Association. 2015. State of the Air 2015. Technical Report. American Lung Association.
[2] Michelle L. Bell, Keita Ebisu, and Roger D. Peng. 2011. Community-level spatial heterogeneity of chemical constituent levels of fine particulates and implications for epidemiological research. Journal of Exposure Science and Environmental Epidemiology 21, 4 (07 2011), 372–384. http://dx.doi.org/10.1038/jes.2010.24
[3] Aspen Hopkins, Pascal Goffin, Miriah Meyer. 2017. Particulates Matter: Assessing Needs for Air Quality Visualization. Posters of the IEEE Conference on Information Visualization (InfoVis). Los Alamitos, CA, USA: IEEE.
[4] Yen-Chia Hsu, Paul Dille, Jennifer Cross, Beatrice Dias, Randy Sargent, and Illah Nourbakhsh. 2017. Community-empowered air quality monitoring system. In Proceedings of the ACM Conference on Human Factors in Computing Systems. ACM, 1607–1619.
[5] Johanna Lepeule, Francine Laden, Douglas Dockery, and Joel Schwartz. Chronic exposure to fine particles and mortality: An extended follow-up of the Harvard six cities study from 1974 to 2009. Environmental Health Perspectives 120, 7 (2012), 965–970. https://doi.org/10.1289/ehp.1104660
[6] PurpleAir. PurpleAir Air Quality Monitoring: An air quality monitoring network built on a new generation of "Internet of Things" sensors (http://map.purpleair.org/). Accessed 2018-08-15. (2017). https://www.purpleair.com/
[7] Liuhua Shi, Antonella Zanobetti, Itai Kloog, Brent A. Coull, Petros Koutrakis, Steven J. Melly, and Joel D. Schwartz. Low-Concentration PM(2.5) and Mortality: Estimating Acute and Chronic Effects in a Population-Based Study. Environmental Health Perspectives 124, 1 (01 2016), 46–52. https://doi.org/10.1289/ehp.1409111
[8] Susanne Steinle, Stefan Reis, and Clive E. Sabel. 2013. Quantifying human exposure to air pollution–moving from static monitoring to spatio-temporally resolved personal exposure assessment. The Science of the Total Environment 443 (01 2013), 184–93. https://doi.org/10.1016/j.scitotenv.2012.10.098
[9] World Health Organization. Ambient air pollution: A global assessment of exposure and burden of disease, 2016

Damla Çay[1], Asım Evren Yantaç[1], Till Nagel[2]
[1] Koç University, [2] Mannheim University of Applied Sciences

WHAT WE LEARNED THROUGH USER-CENTERED EVALUATION OF GEOVISUALIZATIONS

1. RESEARCH PURPOSE

Similar to any other design domain, evaluation is an essential phase in the usercentered design process of geovisualization tools. Different types of evaluation methods can be insightful for different phases of the design process. In this study, we share lessons learned about the evaluation methods employed during the design process of an urban data visualization that facilitates citizens' decisions on urban socio-cultural events.

2. BACKGROUND

There is a great body of knowledge for evaluating information visualizations [1, 2]. By user-centered evaluation, we mean evaluating the higher levels of Munzner's Nested Model [3], which are domain problem characterization, data abstraction design, and interaction design. In a task-based evaluative methodology, participants complete a set of predefined tasks. More specifically, Andrienko et al. [4] developed task formation methods for geovisualizations. As an alternative or a supportive extension to task-based methodology, North [5] suggested the insight based methodology. In the insight-based methodology, participants are asked to generate every possible insight as they are exploring a visualization.

A wide array of methods exist to gather quantitative insights from users. There are close-ended questions to evaluate tasks after a user finishes a task [6] or questionnaires about users' exploration experiences [7] or the usability of a system [8]. In another usability testing method, the think-aloud protocol proposed by Lewis et al. [9], users think aloud when performing the tasks to record and understand the inner decisions or questions participants have. Qualitative methods like semi-structured or structured interviews are also commonly used when evaluating information visualizations as they generate more detailed and rich feedback.

3. METHODOLOGY

Design process started with outlining the information needs using focus group study and participatory design workshop [10, 11]. Based on the needs, we then explored different visualization styles to create a useful tool. To gather feedback on these different visualization styles, and to understand which works best for the defined tasks, we developed a prototype visualizing the same data in the two different manners and conducted a comparative user study of two different visualization styles of the same data. Having two full separate visualization styles is hard to compare in a lab study. But we wanted to get initial qualitative user feedback of the two integrated, rich visualization styles. We used think-aloud, task-based usability testing with 12 participants. After the test, participants evaluated the visualization styles using an item proposed by Hearst et al. [12]. We also employed SUS Usability Scale to measure the general usability. The user study ended with a short semi-structured interview to collect qualitative feedback.

After the user study, we progressed by making some design iterations. However, some evaluation methods were ineffective and we repeated the user study with a different evaluation strategy. In the first study, there was not a significant difference among different visualization styles while in the second study, one visualization was preferred over the other by all participants. We believe some differences in results were related more about the evaluation methods than the visualizations. In our revised evaluative study, we used a combination of a task-based and insight-based methodology with 16 participants. We used both methods together since using only a task-based method might cause participants to focus more on the interface and the general functionality of the tool rather than the visualizations, which was the main focus of the study. After the test, participants evaluated two visualization themes with a survey [7] that aims to measure the positive and negative aspects of their exploration experience. The study was finished with a short semi-structured interview similar to our previous version.

4. IMPLICATIONS

In the first study, the nature of our predefined task-based user study disturbed the casual exploration flow of the geovisualizations. While this allowed us to better measure participants' performance and error rate better, some participants seemed to be constrained by these tasks. For some participants, it was challenging to keep in mind some of the complex task goals. This caused them to end tasks incomplete, which in return reduced the success rate. Others felt the tasks to be too far from their own personal desires. Overall, having such set tasks may hinder the iterative refinement of a data visualization to be used in more casual settings.

We observed that task-based study helps revealing interface problems while insightbased-study generated rich insights on what the visualization affords visually. Based on our experiences, we suggest defining the most important objectives and letting users create their own tasks. In addition to the background questionnaire at the beginning of the think-aloud study, a set of questions to understand the participants' choices regarding the context (what type of an experience they prefer when going out for an event, in our case) were effective to create a more relatable experience.

When comparing visualization styles, some participants from the first study selected more commonly used methods over novel ones and justified their decisions by saying they are more familiar to them. This may be a result of familiarity heuristic [13], which happens when the familiar is favored over the novel. We suggest using an insight based methodology to diminish this effect, as generating insights enabled users to experience the visualizations more thoroughly. Quantitative measures were used in both sessions, gave an overall idea about users' tendencies, however when doing a comparison, generally, there were not any significant differences.

Among the different types of visualizations, geovisualizations are one of the most commonly used ones in everyday life. Our participants were all using map tools on a daily or weekly basis. This widespread use creates certain advantages and difficulties for evaluating new geovisualizations. Familiarity can be an advantage as users have a frame of reference for map tools and they can adjust to different geovisualisations easily. However, it can also be a difficulty. Design patterns create expectations about the visualizations and interface, and if a novel geovisualization does not fit these expectations, users can overlook its advantages easily.

5. IMPACT

Geovisualization tools for everyday life are becoming widespread. Even though there are many visualization evaluation techniques, there is a significant need for refining these methods for everyday use cases. Our insights into evaluation methods can be beneficial for researchers and designers of geovisualization tools.

REFERENCES

[1] E. Bertini, A. Perer, C. Plaisant, and G. Santucci, BELIV'08: Beyond time and errors: novel evaluation methods for information visualization. ACM, 2008.
[2] L. C. Koh, A. Slingsby, J. Dykes, and T. S. Kam, "Developing and applying a usercentered model for the design and implementation of information visualization tools," in Information Visualisation (IV), 2011 15th International Conference on, pp. 90–95, IEEE, 2011.
[3] T. Munzner, "A nested process model for visualization design and validation," IEEE Transactions on Visualization & Computer Graphics, no. 6, pp. 921–928, 2009.
[4] N. Andrienko, G. Andrienko, and P. Gatalsky, "Exploratory spatio-temporal visualization: an analytical review," Journal of Visual Languages & Computing, vol. 14, no. 6, pp. 503–541, 2003.
[5] C. North, "Toward measuring visualization insight," IEEE computer graphics and applications, vol. 26, no. 3, pp. 6–9, 2006.
[6] N. Andrienko, G. Andrienko, H. Voss, F. Bernardo, J. Hipolito, and U. Kretchmer, "Testing the usability of interactive maps in commongis," Cartography and Geographic Information Science, vol. 29, no. 4, pp. 325–342, 2002.
[7] M. A. Yalçin, N. Elmqvist, and B. B. Bederson, "Cognitive stages in visual data exploration," in Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization, pp. 86–95, ACM, 2016.
[8] J. Brooke et al., "Sus-a quick and dirty usability scale," Usability evaluation in industry, vol. 189, no. 194, pp. 4–7, 1996.

[9] C. Lewis, “Using the’thinking-aloud’method in cognitive interface design,” Research Report RC9265, IBM TJ Watson Research Center, 1982.
[10] Ö. Genç, D. Çay, and A. E. Yantaç, “Participatory explorations on a location based urban information system,” in International Conference of Design, User Experience, and Usability, pp. 357–367, Springer, 2015.
[11] D. Çay and A. E. Yantaç, “Learning from the users for spatio-temporal data visualization explorations on social events,” in International Conference of Design, User Experience, and Usability, pp. 124–135, Springer, 2016.
[12] M. A. Hearst, P. Laskowski, and L. Silva, “Evaluating information visualization via the interplay of heuristic evaluation and question-based scoring,” in Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, pp. 5028–5033, ACM, 2016.
[13] C. W. Park and V. P. Lessig, “Familiarity and its impact on consumer decision biases and heuristics,” Journal of consumer research, vol. 8, no. 2, pp. 223–230, 1981.

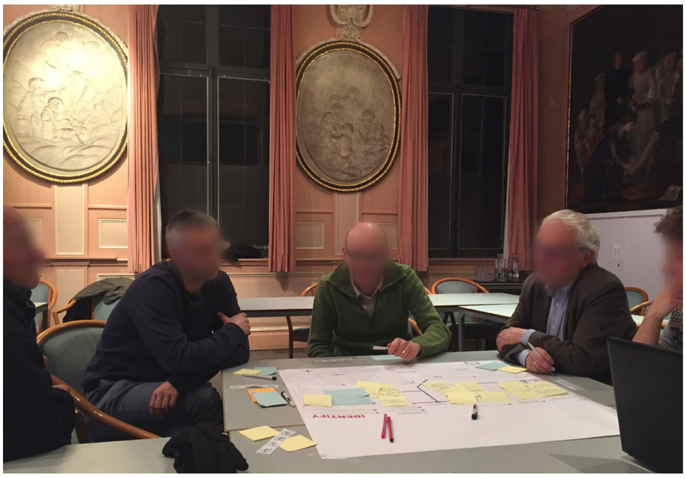


Figure 1: Phase 1 of the CDK methodology: workshop with engaged citizens identifying a target audience, location, context and data (sources).

The outcome of the workshops deployed in the city park. "Are you concerned about air quality in Leuven? Vote here!", in Dutch.

Jorgos Coenen[1], Maarten Houben[1,2], Andrew Vande Moere[1]
[1]Research[x]Design – Department of Architecture, KU Leuven, Belgium
[2]Industrial Design Systemic Change Group, Eindhoven University of Technology, the Netherlands
{firstname.lastname}@kuleuven.be

CITIZEN DIALOGUE KIT:
THE SITUATED VISUALIZATION OF OPEN AND CITIZEN SCIENCE DATA ON PUBLIC DISPLAYS

ABSTRACT
In this paper we introduce our ongoing research on how to leverage the visualization of open and citizen science data within public space to inform and engage citizens. We developed an open-source toolkit, coined "Citizen Dialogue Kit" that is able to convey timely data visualizations on a set of interactive, wirelessly networked displays that can be freely positioned in urban space. The toolkit consists of a participative methodology to help guide the choice of the data and the design of its visualization, a set of off-the-shelf hardware components, and custom-made open source software that controls the whole system. We summarize the design of the toolkit and its initial deployment, and conclude by discussing implications for urban visualization.

Keywords: urban visualization, public displays, public visualization, situated visualization, civic technology, smart cities, participation

1. INTRODUCTION AND PURPOSE

Due to how public decisions are increasingly initiated from bottom-up demands yet often grounded in data-driven evidence, citizens are becoming active data collectors to advocate for local changes. Popular tools such as the Smart Citizen Kit [1] or Air Quality Egg [2] have opened the door for citizens to gather data to hopefully support the, often hyperlocal, causes they care about. However, often relevant data that might support their goals is stored on online platforms, failing to reach, inform or engage those in the community. As such, data that aims to capture the situation of a space, is almost never fed back to those that can influence that situation. While several toolkits exist to collect data and use citizen science techniques to build communities for change [3], none seem to support the situated visualization of the collected data.

In this paper, we describe our ongoing work to develop and evaluate a toolkit for situated urban visualization [4] aimed at conveying citizen science and open data in context.

The "Citizen Dialogue Kit" (CDK) has two objectives: (1) to enable organizations and bottom-up initiatives to present locally relevant data within its physical and sociocultural context, and (2) to enable passers-by to (re)act upon this information and start dialogs. We believe data is most relevant in its context and aim for anyone to be able to encounter it in opportunistic ways without common technical barriers such as QR codes or website URLs. Previous research has demonstrated that situated visualization [5] can broaden and facilitate participation [6] and that contextually located displays can facilitate knowledge sharing in a community [7]. The toolkit includes an interactive polling functionality alongside the urban visualization to potentially engage onlookers [8] and collect their opinions that might be based on the displayed data and thus become part of the dialog.

2. BACKGROUND

Public displays and situated technologies enable a potential to inform and engage citizens about hyperlocal issues through representations of open data from civic platforms and citizen science [1], as well as through interactive polling [8-10]. It offers an opportunity to underline personal links to local issues [11]. Our previous research on urban visualization [4] has already highlighted the use of narrative techniques in visualization [12] to increase personal relevance and aide insight generation on public displays [13]. We have emphasized the impact of contextual clues that aide but also influence and steer insights [14], for example through tacit knowledge and implied social and political values and assumptions [15]. Most recently we have reported on how small displays that present local data on the façades of households are interpreted differently by passers-by based on the social relations, connotation and reputations of the households [16].

3. CITIZEN DIALOGUE KIT

The toolkit was developed over a period of 2 years. During the first year we crystalized its concept, through a series of iterative adjustments based on the results from case studies with low fidelity prototypes in three different European cities (Santander, Antwerp, Aarhus) [16]. In the second year, we developed more robust prototypes, released them as open source [17], and developed a participative methodology that aims to aid citizens and other organizations in deploying the toolkit effectively. Similarly, to the initial concept, the methodology was refined through the critical analysis of iterative workshops with municipal workers and citizens in two European cities (Aarhus and Leuven).

3.1 CDK Methodology

The CDK methodology consists of two workshops (phase 1-2), technical instructions (phase 3) and guidelines to monitor, maintain and interpret the results of the deployments (phase 4). In the first workshop a topic, target audience, physical context and goal is determined along with the appropriate data and data sources. In the following workshop, participants start from this data to ideate visualization concepts and polling questions. Here, we particularly aim to leverage the use of data visualization to potentially lead to meaningful insights that reach well beyond conveying just facts or (sensor) measurements. Finally, participants mock-up the locations in the neighbourhood where they think the target audience can be reached, as well as to determine a contextually relevant narrative that binds all the visualizations and polling questions.

3.2 CDK Technology

The open source software for the visualization devices is designed to run on off-the-shelf hardware: a 7.5” e-ink display, a 3G-enabled microcontroller and a battery. Together these components allow for flexible and self-reliant deployment. Integrated polling buttons encourage passers-by to reflect and leave their opinion behind (see Figure 3).

4. INITIAL DEPLOYMENT

The local government supported a group of citizens (Civic Lab Leuven) in their bottom-up quest to measure the air quality in the whole city. As such, local citizens were invited to apply together with neighbors to subsidize the acquisition of an air quality sensor that can be mounted on their facades. In participation with these citizens, we offered our toolkit to help determine the most relevant data variables, the most approachable visualization techniques and their potential location within the city.

4.1 Co-creation workshops

We led two workshops with citizens (4 and 6 respectively) engaged in this bottom-up initiative (see Figure 1) to evaluate the first two phases of the CDK methodology.
Sparked by a recent news report on the impact of bad air on young children, the workshop participants quickly came to the agreement that those most sensitive to bad air should be the target audience; young children and their parents. The city park was suggested as a potential context where they might encounter these displays. The goal the participants established was to communicate the sources of unhealthy air and encourage citizens to actively change

behaviors. Polling questions such as “How do you estimate the air quality in Leuven?” or “What are you willing to do to help improve air quality?” were raised for the polling feature. The participants drafted corresponding answering options specifically to relate to everyday life and a personal impact on air quality, e.g. not using coal at a barbecue, no bonfire at camp or a tax on wood stoves. A focus on fire related air pollution emerged from the sessions as a result of the specific sensitivity of the sensors used by the participants for this type of pollutants.
Participants listed the required data sources, including their own measurements, as well as exposure limits set by the WHO and statistics on the sources that contribute the unhealthy air. In the second workshop the visualization of this data was conceptualized with the goal of easy readability, including by children. A simple line chart would communicate their air quality measurements from the past few days with a clear indication of exposure limits. Pictorial charts were suggested to visualize the sources of unhealthy air. The narrative proposed by the participants leads the passer-by from the air quality measurements over to the sources of pollution and finally to the poll about the actions they could take as a conclusion and topic for debate.

4.2 Deployment

Based on the outcome of these workshops we subsequently deployed the wireless displays in the city park over a period of several days to get a first impression of their performance in a real-world urban environment (see Figure 2-3). Although it remains challenging to draw the attention of passers-by [18], people who frequently pass by the location seem to notice something new and unusual, often mentioning that the colors drew their attention and that they were convinced to stop and take a closer look when seeing the embedded electronics (“Not just an advert, something interactive”). Regularly moving the installation to trigger this novelty effect may be a strategy to increase visibility. Initial observations and interviews point to increased engagement when polling is enabled (e.g. “Hey, you can’t vote yet, you need to read the chart first!”). The deployment also indicated a potential to trigger reflections that relate the data to the local context (e.g. about a peak in pollution, “Why then? Does it have to do with less students in the city?” or “It will get worse here, soon the busses will be rerouted”).

5. IMPLICATIONS

The first results of the deployment of our CDK revealed several preliminary implications for the design of situated visualization.

Figure 3: A passer-by leaving his opinion on local air quality levels.

Visualization literacy and comprehension. Public visualizations [4] generally aim for a wide audience. This implies varying levels of visualization literacy and expertise on the topic addressed. Through narratives we can aid comprehension [13], but we have observed that if understanding is lost, passers-by did not participate in the poll. For instance, one passer-by said “Is one hundred good or bad? I don’t know so I didn’t vote”. Another explained “I didn’t vote because I don’t know what PM means”, subsequently he did vote on the second display, saying that the “pictograms were easy to understand and I want to contribute”.
Contextual relevance. While the extreme flexibility and mobility afforded by the wireless 3G connection and battery enables a potential for a strong contextual and meaningful relevance of the visualization to its environment, the physical (e.g. space syntax), temporal (e.g. fluctuating air quality) and social (e.g. demographic) implications should equally be considered in each context.
Resolution and monochrome. The e-ink display technology, chosen for its low energy usage and resulting portability, only allows a small display resolution that limits the complexity and sophistication of the data visualization. Combined with the monochrome style of rendering, the potential visual encodings are limited. We therefore looked to older seminal works such as Bertin for inspiration [19].
Data update rate. The update rate of the visualized data can be determined based the data source, context and the desired duration of the deployment. While a lower data update frequency is beneficial to battery life, it can lead to a disconnect within the user interaction flow (e.g. when the next update is expected). However, it has also been suggest-

ed that a low update rate could perhaps focus attention to specific moments through scarcity [20].

Display refresh rate. The e-ink display technology has a slow refresh rate (larger than 6 seconds), requiring similar consideration when designing the interactive modalities (e.g. for polling) or narratives (e.g. choosing multiple displays over pagination on a single one). We used a buzzer as an auditory confirmation of a vote, but have also considered integrating LEDs.

6. CONCLUSION

We have outlined our ongoing research into the bottom-up creation of interactive situated visualizations, i.e. representing local data within the public environment to which passers-by can react. We summarily described the design and first deployment of the tool and discussed some implications for designing situated urban visualization. In future work, we will continue to refine the CDK methodology, explore the design space of the technology and evaluate the impact on participants, passers-by and the wider neighborhood.

ACKNOWLEDGEMENTS

This project has received funding from OrganiCity.eu, a project in the European Union's Horizon 2020 research and innovation program under the grant agreement No. 645198. We also wish to thank Civic Lab Leuven and LeuvenAir.be for their participation.

REFERENCES

[1] "Smart Citizen Kit." [Online]. Available: https://smartcitizen.me. [Accessed: 05-Jul-2018].
[2] "Air Quality Egg." [Online]. Available: https://airqualityegg.wickeddevice.com. [Accessed: 05-Jul-2018].
[3] "Making Sense." [Online]. Available: http://making-sense.eu. [Accessed: 05-Jul-2018].
[4] A. Vande Moere and D. Hill, "Designing for the Situated and Public Visualization of Urban Data," J. Urban Technol., vol. 19, no. 2, pp. 25–46, 2012.
[5] W. Willett, Y. Jansen, and P. Dragicevic, "Embedded Data Representations," IEEE Trans. Vis. Comput. Graph., vol. 23, no. 1, pp. 461–470, Jan. 2017.
[6] S. Lindley, A. Thieme, A. Taylor, V. Vlachokyriakos, T. Regan, and D. Sweeney, "Surfacing Small Worlds through Data-In-Place" Comput. Support. Coop. Work, Feb. 2017.
[7] C. Liu et al., "Pinsight: A Novel Way of Creating and Sharing Digital Content Through 'Things' in the Wild," in Proceedings of the 2018 Designing Interactive Systems Conference, New York, NY, USA, 2018, pp. 1169–1181.
[8] L. Koeman, V. Kalnikaité, and Y. Rogers, "'Everyone Is Talking About It!': A Distributed Approach to Urban Voting Technology and Visualisations," in Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, New York, NY, USA, 2015, pp. 3127–3136.
[9] L. Hespanhol, M. Tomitsch, I. McArthur, J. Fredericks, R. Schroeter, and M. Foth, "Vote As You Go: Blending Interfaces for Community Engagement into the Urban Space," in Proceedings of 7th International Conference on Communities and Technologies, 2015, pp. 29–37.
[10] N. Valkanova, S. Jorda, M. Tomitsch, and A. Vande Moere, "Reveal-it!: The Impact of a Social Visualization Projection on Public Awareness and Discourse," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2013, pp. 3461–3470.
[11] A. Taylor et al., Data-in-Place: Thinking Through the Relations Between Data and Community. New York, NY, USA: ACM – Association for Computing Machinery, 2015.
[12] E. Segel and J. Heer, "Narrative Visualization: Telling Stories with Data," IEEE Trans. Vis. Comput. Graph., vol. 16, no. 6, pp. 1139–1148, Nov. 2010.
[13] S. Claes and A. Vande Moere, "The Impact of a Narrative Design Strategy for Information Visualization on Public Display," in DIS'17, 2017.
[14] A. V. Moere and D. Offenhuber, "Beyond ambient display: a contextual taxonomy of alternative information display," Int. J. Ambient Comput. Intell. IJACI, vol. 1, no. 2, pp. 39–46, 2009.
[15] A. Vande Moere and N. Wouters, "The Role of Context in Media Architecture," in Proceedings of the International Symposium on Pervasive Displays, 2012, pp. 12:1–12:6.
[16] S. Claes, J. Coenen, and A. Vande Moere, "How Context Impacts Engagement with Spatially Distributed Public Visualization and Polling Displays," in Proceedings of the 2018 NordiCHI conference, Accepted for publication.
[17] "Citizen Dialogue Kit," Github. [Online]. Available: https://github.com/citizendialogkitproject/citizen-dialog-kit. [Accessed: 23-Jul-2018].
[18] J. Müller et al., "Display Blindness: The Effect of Expectations on Attention Towards Digital Signage," in Proceedings of the 7th International Conference on Pervasive Computing, Berlin, Heidelberg, 2009, pp. 1–8.
[19] J. Bertin, "Semiology of graphics: diagrams, networks, maps," 1983.
[20] N. Davies, S. Clinch, M. Mikusz, O. Bates, H. Turner, and A. Friday, "Better off: When Should Pervasive Displays Be Powered Down?," in Proceedings of the International Symposium on Pervasive Displays, 2017, pp. 19:1–19:9.

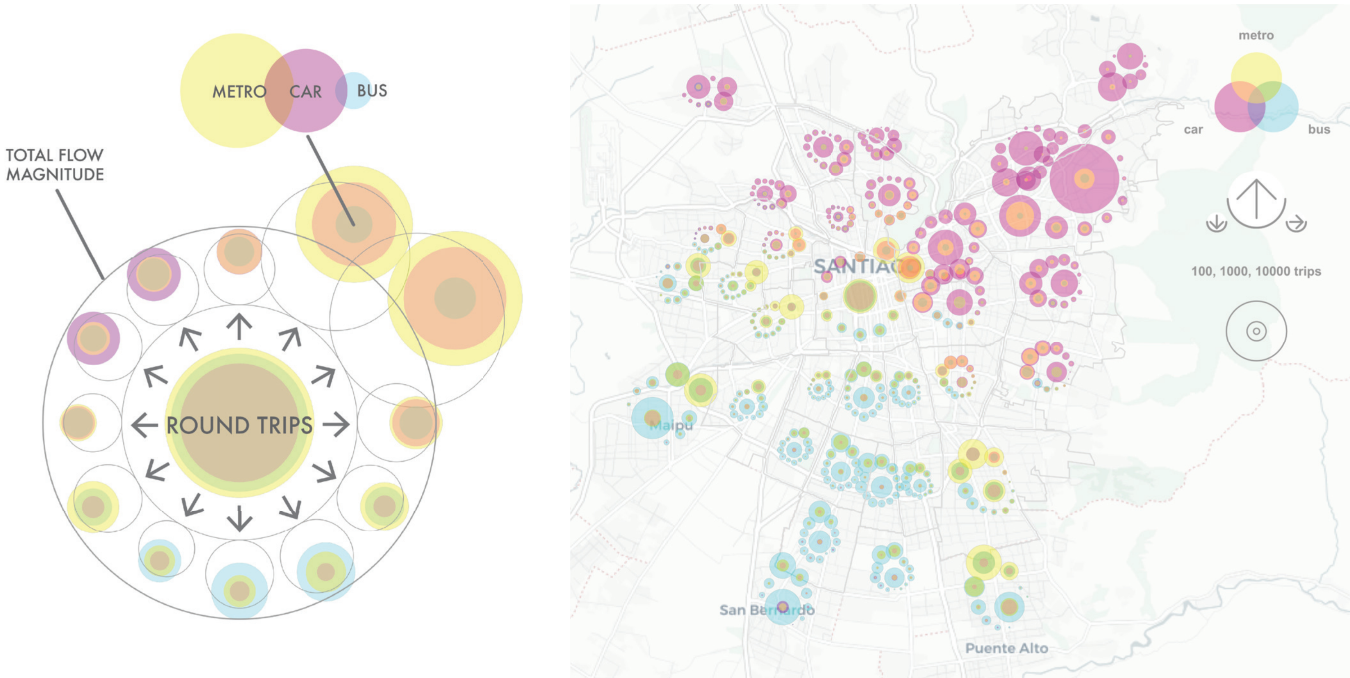


Fig. 1. Left: Glyph with schema. Empty circles represent direction-wise flow magnitude, with round trips at the center. Bolder line is total flow. Each colored circle results from the superposition of the flows disaggregated by transport mode. Right: Overview of commuting flows in Santiago, Chile, at the morning peak. Clear patterns immediately emerge: predominance of certain modes of transportation over different zones, differentiation of transport mode by destination and a monocentric city structure.

Ignacio Perez-Messina[1] & Eduardo Graells-Garrido[1]
[1] Data Science Institute, Faculty of Engineering, Universidad del Desarrollo, Santiago, Chile
{iperezm,egraells}@udd.cl

TOWARD VISUALIZING COMMUTING PATTERNS FROM MOBILE PHONE DATA

Keywords: Glyph-based Visualization; Urban Informatics; Commuting; Flow Data; Mobile Phone Data.

1. RESEARCH PURPOSE AND BACKGROUND

The availability of large amounts of data has allowed to study urban phenomena at previously unseen scales [1]. One of these sources is the set of billing records from mobile phone networks, known as Data Detail Records (XDR). XDR provides a cost-effective way to analyze city-related data, because mobile operators already generate and store it. In this project we focus on commuting patterns inferred from XDR [2], as understanding commuting flows can help to plan and manage a city's transportation network. In our data, in addition to origin and destination, each commuting trip has one or more modes of transportation (e.g., metro and bus), and potentially some within-trip waypoints, which are used to infer the modes of transportation using work-in-progress machine learning methods [3]. Note that, in contrast with GPS data used in similar contexts, XDR has a coarser spatial resolution (the position of each waypoint is the corresponding mobile phone tower), and a sparser temporal resolution (there is an XDR record every 15 minutes). Analyzing and visualizing this information presents several challenges, from the visual abstraction of multivariate data, to the spatio-temporal granularity that makes known approaches impractical. Telcovis [4], a prominent visualization for XDR, offers a complex system for analyzing this data but lacks integration and transport mode disaggregation.

2. APPROACH

Our aim is to create an interactive visualization of commuting patterns to perform transportation analysis tasks with a more versatile glyph-based approach than current ones [5,6,7].

Before designing our first prototype, we identified relevant tasks for domain experts in transportation. We found three tasks: Task 1, to have an overview of flows within a city, having into account departure time and modes of transportation; Task 2, given a zone A of origin or destination (e.g., a neighborhood), visualize the outgoing / incoming / passing flows from / to / through A, to analyze the transportation patterns of a specific population; and Task 3, to compare XDR-inferred commuting patterns with available ground-truth data, such as travel surveys. This last task allows to build trust with domain experts, who are used to these traditional data sources, however, such datasets are expensive and limited in terms of spatial and temporal scope.

We designed a simple glyph to represent flow magnitude by mode of transportation (for up to three different modes, in this case: car, metro and bus), and a super-glyph structure to establish flow direction (Fig. 1, Left). The glyph is composed by three concentric circles. Each circle is associated with a mode of transportation by its color, which is partially transparent, so different modes mix up in a subtractive way. Pure CMY hues were used to associate a circle, whose area is mapped to flow magnitude, to mode of transportation, so the resulting glyph is: an outermost halo of the predominant mode of transportation of a pure CMY color, an inner halo of the pure RGB color resulting from the mix of the predominant and second most predominant modes, and an innermost grayish circle representing the least dominant mode. This way, the predominant mode of transportation is emphasized as it is always the outermost and brightest color, allowing for a pre-attentive spatial sense of predominant modality. The glyphs are then organized into a super-glyph which confers the represented flows a direction (depending on its place within the structure) and assigned to each administrative division of the city (Fig.1, Right), in fulfilment of Tasks 1-2 (flow sense can be changed manually). Task 3 is left for future work, as it involves understanding the priorities of transportation experts when comparing results. We will perform a series of pilot studies to find an approach that can be integrated into our proposed design.

3. IMPLICATIONS AND IMPACT

Our visualization allows for new tasks to be performed and its organic design leaves plenty of room for expansion and opens up new possibilities for mobility analysis. To validate our task and visual abstraction, we are working closely with experts from the transportation authority in Santiago. Moreover, for tools like ours to be used in planning contexts, the source algorithms must be interpretable, a quality that can be supported by detailed visual exploration of algorithmic results.

REFERENCES

[1] Blondel, V. D., Decuyper, A., & Krings, G. (2015). A survey of results on mobile phone datasets analysis. EPJ Data Science, 4(1), 10.

[2] Graells-Garrido, E., & Saez-Trumper, D. (2016, May). A day of your days: estimating individual daily journeys using mobile data to understand urban flow. In Proceedings of the Second International Conference on IoT in Urban Space (pp. 1-7). ACM.

[3] Graells-Garrido, E., Caro, D., & Parra, D. (2018). Toward Finding Latent Cities with Non-Negative Matrix Factorization. In: Said, A., & Komatsu, T. (eds.) Workshop on User Interfaces for Spatial-Temporal Data Analysis (2018). http://ceur-ws.org/Vol-2068/uistda4.pdf

[4] Wu, W., et al. (2016). Telcovis: Visual exploration of co-occurrence in urban human mobility based on telco data. IEEE transactions on visualization and computer graphics, 22(1), 935-944.

[5] Andrienko, G., et al. (2017). Revealing patterns and trends of mass mobility through spatial and temporal abstraction of origin-destination movement data. IEEE TVCG, 23(9), 2120-2136.

[6] Ma, Y., et al. (2016). Mobility viewer: An Eulerian approach for studying urban crowd flow. IEEE Transactions on Intelligent Transportation Systems, 17(9), 2627-2636.

[7] Zeng, W., et al. (2013, June). Visualizing interchange patterns in massive movement data. In Computer Graphics Forum (Vol. 32, No. 3pt3, pp. 271-280). Oxford, UK: Blackwell Publishing Ltd.

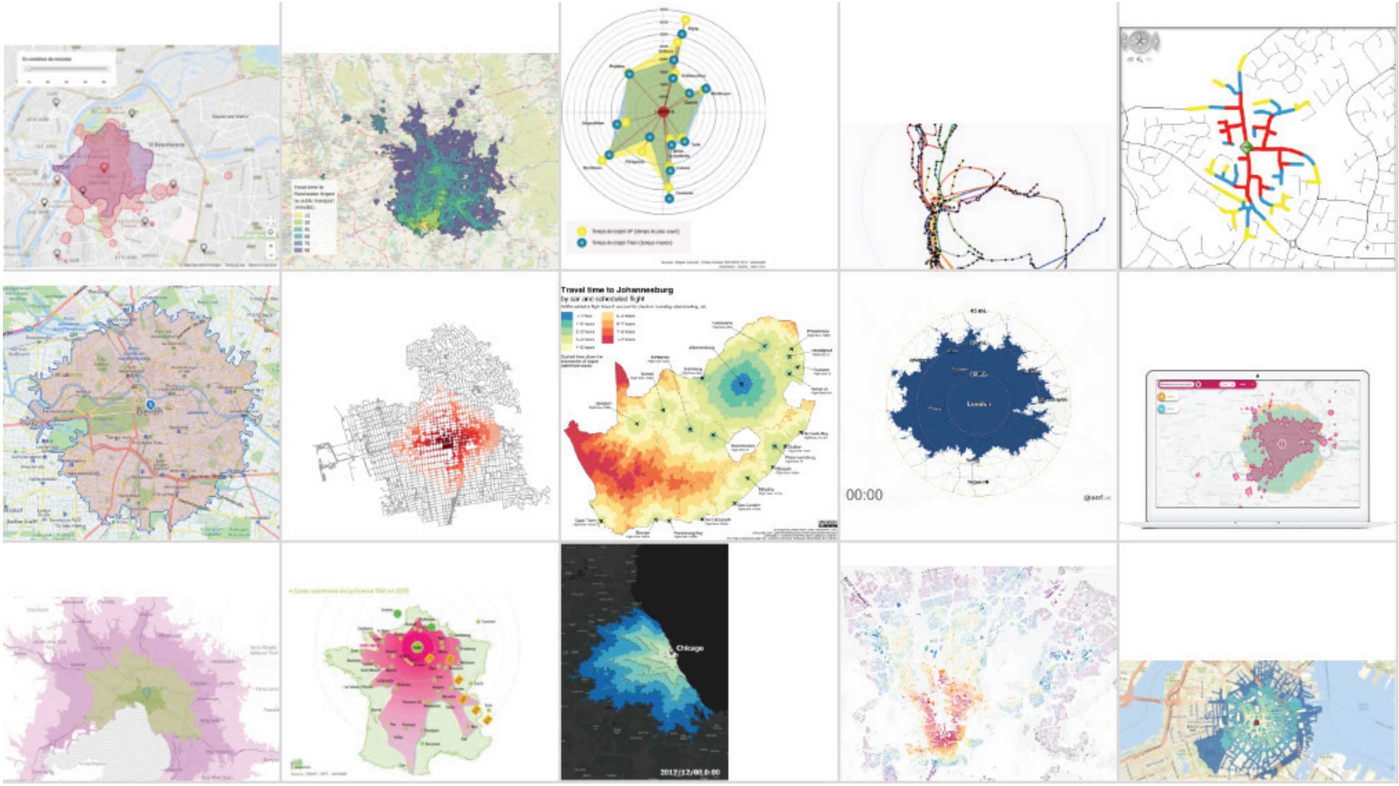


Figure 1: Sample of isochrone maps we collected and organized in our corpus.

Romain Vuillemot[1], Thomas Leysens[1], Philippe Rivière[1], Aurelien Tabard[1]
[1] LIRIS CNRS, Univ. Lyon, France of Technology, the Netherlands

AN ONLINE CORPUS OF ISOCHRONE MAPS

Keywords: Isochrones Maps, Design Space, Corpus

1. CONTEXT AND CONTRIBUTION

Isochrone maps depict destinations that can be reached within a time range (e.g. 5min). Such maps are popular in GIS tools and on mobility websites as they are intuitive for non-technical people. Using isochrones, urban decision-makers can visually capture the coverage of transit systems, and spot difficult to reach areas. City inhabitants can use them to find living areas that are close to work, daycares, or other personal preferences.

We present a web-based corpus of isochrones maps we collected, to identify their design space. While isochrones are constantly improved to support new tasks (e.g. to display time-varying mobility data [1], or support multi-modal networks characteristics [2]) the breadth of their design space (e.g. color scales, color gradients, steps, data types) still needs to be better understood. We built this corpus from diverse sources including Google Image, GIS websites, isochrones toolkits galleries, digital libraries, and personal collections. We gathered 52 isochrone maps (and counting). We then sanitized the dataset by removing redundant representations and un-related ones, to keep 37 representative isochrone maps.

The corpus is available online as a faceted search website[1]. The website allows filtering by properties such as color scale, color interpolation or any other metadata we identified (e.g. year, data source). In order to improve the

breadth of the corpus, we turned it public, the community can submit structured examples using a Google Form[2]. And the corpus is available as an open spreadsheet[3], that can be exported and further analyzed, or simply commented to improve quality.
As we already started to perform quantitative analysis on the corpus, we also provide all the code that allow to extract statistics (e.g. parameters frequency) from the spreadsheet using a Python Notebook[4] to further update our analysis.

We plan to continuously update the corpus to reflect both past and present advances in isochrones designs and applications. Such corpus subscribes to other community-based geo-related collections such as the Visualizing Cities Open Platform website[5]. It will also be used in the frame of the M2I project[6] that partially funds this research, and which aims at providing new visual tools to to build and follow combined itineraries in urban areas.

2. PRELIMINARY REPORT

We now report on a very preliminary analysis of our corpus. Figure 1 shows a sample of visuals from our collection. Regarding the isochrones creation period, they range from 1881 to 2018, where the earliest are paper-based representations. Recent isochrones are computer-generated images resulting from the assembly of multiple building blocks (e.g. base map, data source, layers, shapes generator). We haven't (yet) identified any particularly frequent combination of those blocks. This was surprising as, from our experience, isochrones are difficult to generate due to the complex data queries they rely on, so we expected less isochrone generators to exist. However, a common denominator was the data source used: often Navitia or OpenTripPlanner — both are very popular, free itinerary APIs.

Color is a key element of isochrone maps design, as each region usually has a filling encoding the time steps. Yet we found very diverse color palettes and we could not identify any unique strategy from our corpus analysis. Color interpolation was linear for roughly a half of our examples (the Viridis one being very popular), the other half used thresholds or categories. Most of the color scales could arguably be improved (by not using rainbow colors, by being color-safe, etc.).

We expect to continue this analysis and discussion during the CityVis'18 workshop and report on more advanced and statistics-supported findings from our isochrone corpus.

FOOTNOTES

[1] https://romsson.github.io/isochrones/web/isochrones.html
[2] https://goo.gl/forms/fkt7RiVzdsVB81vK2
[3] https://goo.gl/TX1gcp
[4] https://romsson.github.io/isochrones/processing.ipynb
[5] http://cityvis.io/collection.php
[6] https://projet.liris.cnrs.fr/mi2/

REFERENCES

[1] F. Gortana, S. Kaim, M. von Lupin, and T. Nagel, Isoscope-Visualizing temporal mobility variance with isochrone maps," Poster Abstracts of IEEE VIS 2014, 2014.
[2] J. Gamper, M. H. Böhlen, W. Cometti, and M. Innerebner, Defining isochrones in multimodal spatial networks," in CIKM 2011, ACM, 2011.

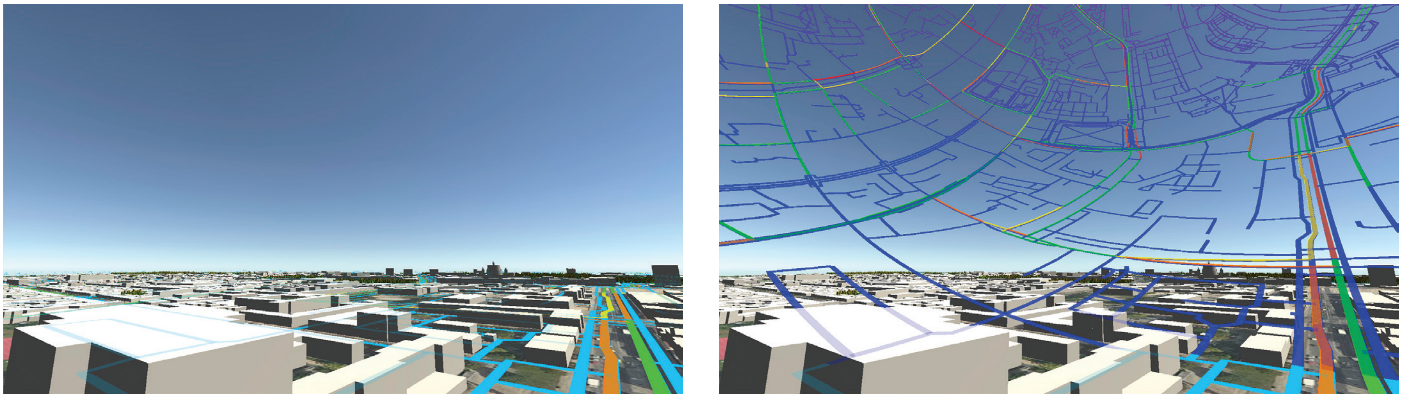


Figure 1: Separating and bending spatial data upwards legibly reveals occluded information in views closer to the ground.

Maxim Spur[1] (maxim.spur@crenau.archi.fr) & Vincent Tourre[1] (vincent.tourre@ec-nantes.fr)
[1] Ecole Centrale de Nantes, France

URBAN DATASPHERE: EXPLORING IMMERSIVE MULTIVIEW VISUALIZATIONS IN CITIES

Keywords: Multiview; Multiperspective; Multilayer; Urban data visualization; Immersive analytics

1. RESEARCH PURPOSE

We investigate and develop possibilities for displaying geospatial data within virtual city models, or over real cities if using AR. The focus here is on providing a viewer immersed in a city's geometry with information on their immediate as well as more distant surroundings while maximizing legibility and discoverability of both without compromising immersion.

2. BACKGROUND

The origin of our problem formulation lies in interviewing professionals and researchers in urbanism about their interactions with urban data and its visualizations to obtain a more complete understanding of place—urban immersion, as well as an extensive literature survey on related topics.

Aside from the trivial solution of having both an overview in form of a top-down map view side-by-side or picture-in-picture with an immersive view of a street or a city block's details, attempts have previously been made to deform the city's geometry to more effectively utilize a viewer's visual field.

3. APPROACH

Instead of deforming the whole city, we propose a method that keeps the urban geometry and therefore its legibility intact by detaching the data layer from the ground plane and only deforming it (Fig. 1).

Our approach is centered around the inverse stereographic projection—a smooth, bijective and conformal mapping that projects points from a plane onto the surface of a tangential or intersecting sphere. The origin of the projection rays lies on the sphere's furthest point from the plane—the end point of the sphere's diameter that is perpendicular to the plane. A line is drawn from this origin to points on the plane, which are projected on the sphere where this line intersects its surface.

This projection can in effect map a plane of arbitrarily large dimensions onto a sphere. The distortions in the projection increase with the distance between original and projected point, resulting in zero distortion in points at which the plane intersects the sphere.

Applied to our proposal for displaying urban data, the

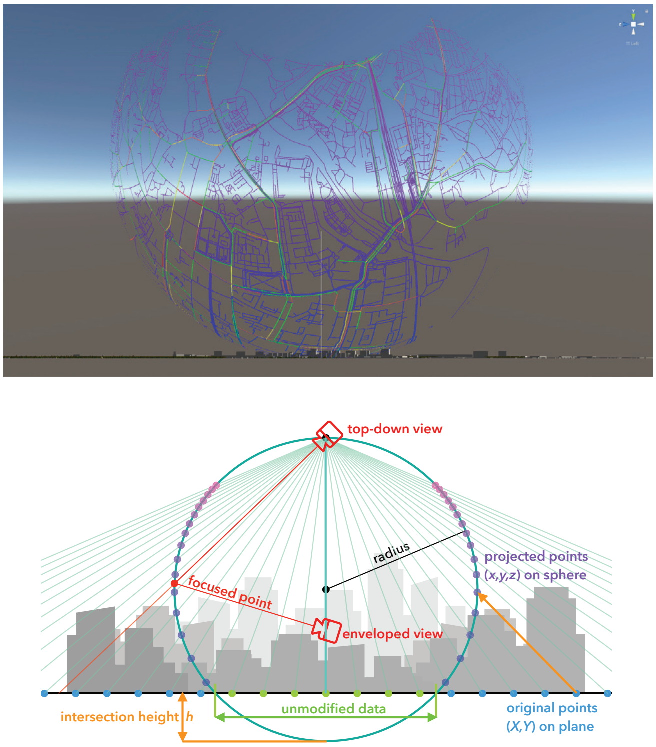


Figure 2: Stereographic projection of geolocated urban data to a sphere enveloping the viewer.

plane represents the city map containing said data (road networks and traffic congestion in our example), with building geometry and textures on top of it (Fig. 2). The projection sphere intersects this plane and contains the camera view. Instead of displaying points that would have been projected below the ground plane, we keep those at the intersection plane at their original position—data that would be legible anyway from the point of view is not distorted. Varying the sphere's radius determines how closely data from further distances is "pulled in" to the visual field (Fig. 3).

As this projection utilizes the vertical visual field to display more distant data, using VR with its potentially much greater field-of-view than traditional displays seems appropriate to take full advantage of this new approach.

Another feature that stereographic projection makes possible is the seamless transition from a flat, top-down view with data at its original position to the enveloped view described above, by moving the point of view to the origin of the projection rays. Starting at this position and smoothly moving down into the enveloped view shows the viewer exactly how the deformation works and where the projected data comes from.

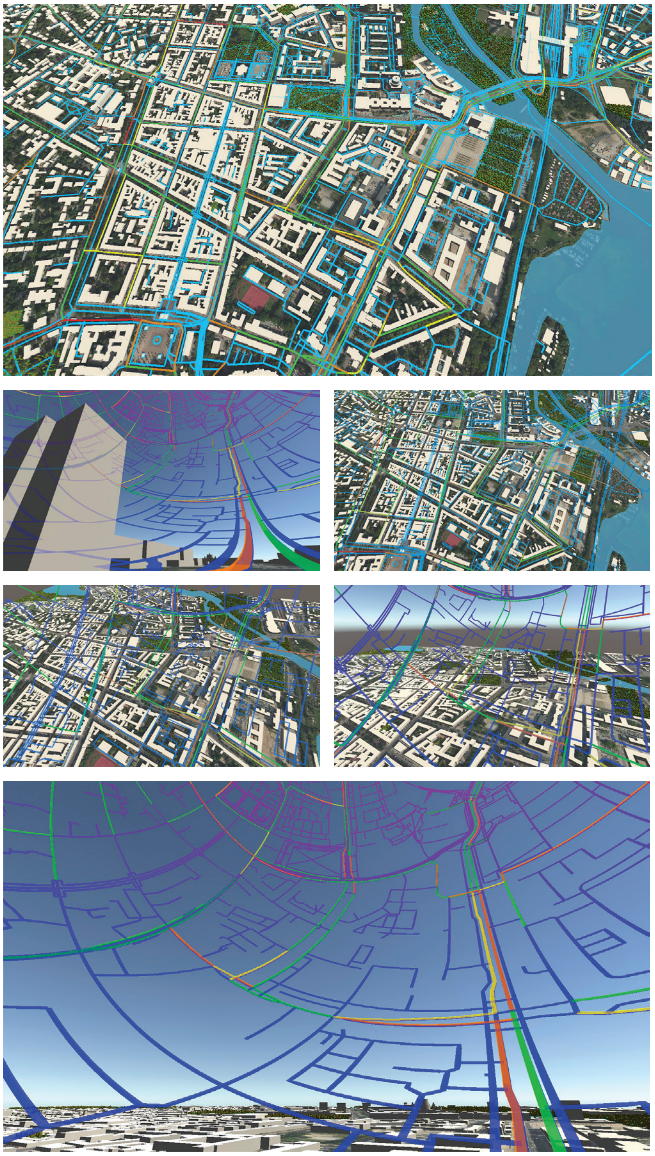


Figure 3: Sequence showing transition from top-down view to en- veloped pedestrian perspective.

4. ORIGINALITY

The novelty of our approach lies in its separation of spatial data from the ground plane—inspired by an approach in geoinformation— and utilizing stereographic projection to display it legibly from a view that would have otherwise rendered it illegible. Compared to the other option of showing data inside an immersive view—via a separate screen, or a (mini-) map—our approach aims to more firmly anchor the data to its spatial position and preserve a better sense of scale. While a traditional map is immediately familiar, the deformation is quickly understood and an understanding of the distances as displayed in the projection sphere can be gained.

By building in the seamless transition from the top-down view into the enveloped perspective and making it a frequent interaction when navigating the map, we help establish an intuitive understanding of how the deformation happens and the relation between projection and original position, which is also something we have not encountered during our literature survey.

5. PRACTICAL IMPLICATIONS

By using MapBox as a source and repository for maps and Unity as the 3D environment, any geocoded data can be easily uploaded and displayed with our projection system. Parameters of the projection (sphere diameter, intersection height) and visualization (colors, transparency, etc.) can be edited on the fly.
Furthermore, Unity allows for porting the system to VR and AR applications for a deep immersion into the data and its context within a virtual or real city.

6. CHALLENGES

Before arriving at the described scheme, we first experimented with different ways of separating and projecting the data on a dome above the viewer, and also with a sphere that is bisected by the ground plane, with the lower half shining through the ground. These two attempts created a disconnect between the data and its original position, e.g. the data closest to the viewer was counter-intuitively projected to distant points like the zenith or straight down.
In refining the current stereographic projection approach, the challenges are finding the optimal sphere radius and intersection height, which strongly depend on the viewer's elevation, the city's geometry and on how far the area of interest extends past the viewer's position. More generally, other challenges are creating useful indicators on how strongly the data is deformed or where it is coming from on the ground plane, and establishing effective navigation schemes, particularly for VR environments.
Finally, this approach's efficacy needs to be experimentally verified to compare it to traditional methods and other projection schemes, like onto a cone or cylinder, as well to find optimal projection and visualization parameters.

7. IMPACT

Based on first rounds of feedback from fellow computer scientist and urbanist researchers in our lab we have encountered great enthusiasm for this system for urban data visualization, along with suggestions for features that we already implemented, such as preserving original data location in the immediate surroundings of the viewer.
Following this feedback, we will conduct experiments to evaluate the system's potential advantages over comparable visualizations in a formal way.
Aside from using this system to display any geolocated urban data such as population density, air/noise pollution, etc. overlaid on road networks, another promising application could be urban navigation, particularly using AR in handheld devices or head-up displays in vehicles.

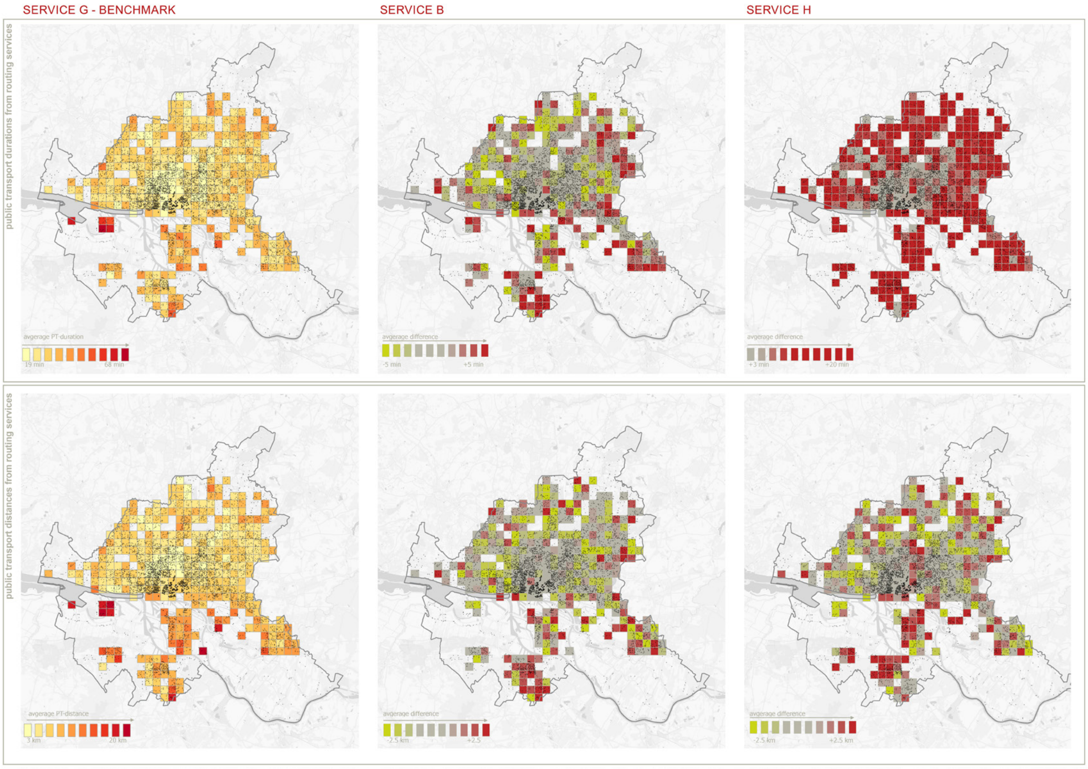


Figure 1: PT travel durations (top row) and distances (bottom row) as calculated by the routing services for Service B and Service H compared to the benchmark

Benno Bock, André Klein[1]
[1] civity - Management Consultants

UNDERSTANDING ROUTING ALGORITHMS

Keywords: Routing, urban mobility, public transport, choice modelling

1. FILTER BUBBLES AND ACCESSIBILITY

How we live together is defined by actions and interactions that are influenced by the options the urban landscape provides us, by how we can move between places and by the information we have on the former two. Information, actions and mobility create a structure we ultimately call city. Recently, a ubiquitous digital layer has gained an enormous influence on this mechanism. A common word used in this context is the “filter bubble” in which an intelligent algorithm influences what a specific person may see or not see. Filter bubbles can also influence the accessibility and connectivity of our cities by pre-selecting the information on how we get from A to B. For example, information on public transport (PT) routes is often derived from routing apps. These services show a list of options in respect to start-time, duration, changeovers, leading and lagging walking distance.

2. ROUTING A REPRESENTATIVE ORIGIN-DESTINATION SAMPLE

As the routing algorithms are usually not open source, the exact mixture is unknown to the user. In this context, civity plans to analyse the algorithms with a combination of API requests and choice modelling. In a first step, a web-mining system for PT-routing services has been set up. As a pre-requisite, a demand-orientated set of origin-destination (OD) choices has been created within

a city – in this paper Hamburg. This set is modelled as relations between raster points and is weighted by the movements for the overlying postal zones which were gained by analysing floating phone data. A large weighted sample of at least 10 000 OD pairs is taken as the basis for the OD requests. As the main focus is PT, our requests to the routing algorithms are limited to those that provide PT routes. We have set the time for the routing to 8:30 for the next workday to represent peak traffic.
We have analysed three of the leading PT-routing services which we will call 'Service B', 'Service G' and 'Service H' in this paper. We consider one to be the benchmark ('Service G'), as it is by far the most popular service of the three services studied here. All of the services are available globally and include real-time information on street traffic velocities and PT departure times. Traffic velocities can influence street-bound PT, such as busses, whereas delayed departures can influence trip durations, for example, when connecting services are missed.
The volume of non-standardised parameters that can be passed to the API increases the complexity of the experiment and has been kept to a standard set which we considered to be comparable over the various services. Uncertainties do remain as it is unclear if PT is routed on a network representing realistic traffic volumes. This is mainly because no traffic parameters have been able to be passed to the API. The work was conducted using web-mining scripts written and adjusted in Python and a PostgreSQL database. The results were initially visualised with QGIS.

3. VARIATIONS FOR INNER CITY TRIPS

For inner city raster cells of 1 km² (defined here as OD relations within the administrative boarders of the City of Hamburg), average durations and distances for outgoing public transport trips were derived. Values were only generated for cells with more than 5 OD relations.
The values indicate that for the benchmark central cells tend to have lower average distances and durations and therefor can be regarded to be better connected. Fringe areas tend to show higher distances and durations. This partly results from the limitation of the scope on inner city relations, but may also be influenced by a lower connectivity by PT-service frequencies and station densities. Additionally, enclaved areas isolated isolated, for instance, by port facilities, harbours, rivers or industrial areas, show the highest deviations in average duration and distance.

In comparison to the benchmark, the average distances of the other two services for the same raster cells vary significantly. The differences seem to spread following a normal distribution varying from -2.5 km for the 0.05-percentile to +2.5 km for the 0.95-percentile with a median of almost 0 km. A similar distribution can be found for the differences of the durations of Service B in regard to the benchmark. Here, the 0.05-percentile lies approximately at -5 min, the 0.95-percentile at +5 min and the median of differences is also approximately at 0 min. The durations for Service H are the only ones with a strong positive bias. Calculated durations for PT routings are unexceptionally longer than the benchmark. Ranging from an average difference of +3 min up to +20 min.
Strong differences between the services are spatially highlighted based on the abovementioned 1 km² raster grid. For Service B, areas with an above-average duration (marked red in Figure 1) seem to alternate with areas with lower durations (marked green in Figure 1) without a specific spatial focus. The only exception here being that central areas tend to have lower absolute differences and some fringe areas to the south and east tend to show lower durations for the benchmark. These areas appear in Figure 1 as red zones with a positive difference for distances and durations. For Service H, raster cells predominantly indicate longer average durations. More central cells tend to have lower positive deviations from the benchmark.

4. PERFORMANCE OF SPECIFIC PT STATIONS

Variations can be seen between services for some specific stations. While the benchmark seems to suggest the main train station (Hamburg Hauptbahnhof) more frequently for rail and light rail service, the other two routing APIs tend to suggest the subway services to and from the main station. Jungfernsteg station, another subway service, also seems to be favoured by these APIs. Other stations with major deviations are the important interchange stations of Altona, Harburg and Berliner Tor (See Fig. 2).

5. Results for Commuting Trips

The routing services also provided a highly varied output for route information on OD relations which we considered to be typical commuting trips. These trips were idealised as trips starting from outside the administrative borders and ending at the town hall in Hamburg city centre. Significant differences were found for both distances and durations. The highest differences were found in

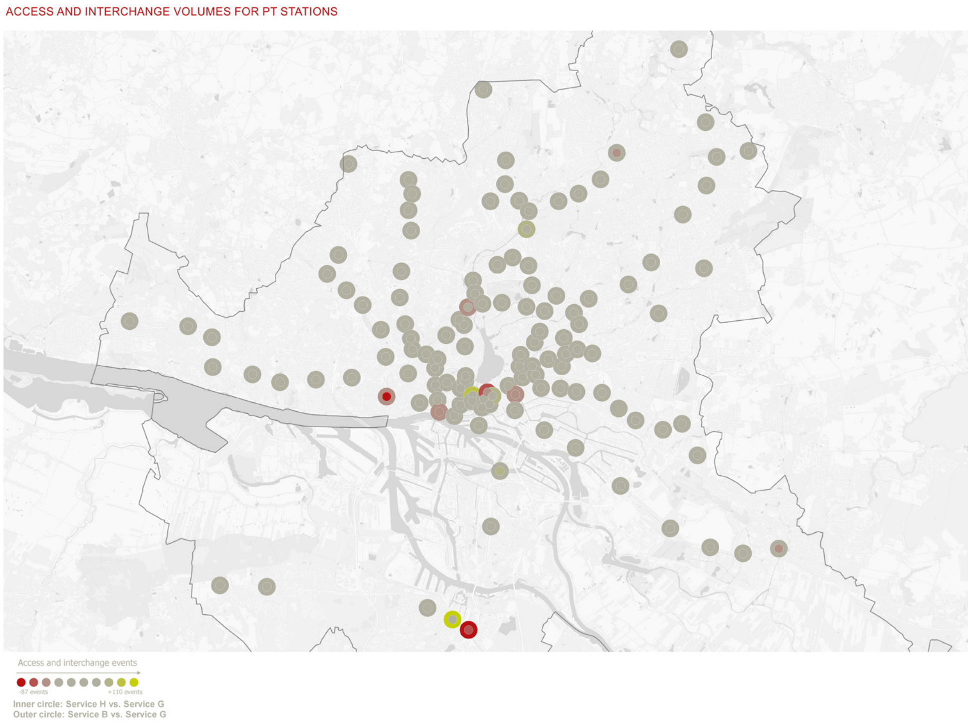


Figure 2: Differences in suggestion volumes for access and interchange and subway, light rail and rail services from Service B and Service H compared to the benchmark

travel times with a variation between -41.7% and 46.8% for Service H in respect to benchmark service G. The variations seem to be strong, to such an extent that some of the services even seem to be strongly misleading.

Interpretation: Potential for Digital-Spatial Segregation
The results show that information on distance and durations of PT trips differs considerably. While the spatial distribution of these differences appears to be fairly random, certain areas show a clustering of cells with either negative or positive variations compared to the benchmark. For most of the outer districts of Bergedorf or Harburg, the results could lead to the conclusion that an exclusive user of Service B may appear to be not as well-connected as an exclusive user of Service G in the same area. The results of Service H indicate that exclusive users of this service may generally feel less connected by PT. Interestingly, some zones show negative differences for durations while at the same time showing positive differences for distances. The results indicate that prioritisation and selection of routes that are communicated to the user of the services may be based on very different principles.

6. FURTHER STEPS

The choice modelling software Biogeme is now to be used for further analysis. In a later step, significant output may be visualised in an interactive storyboard using standard web templates for the Django-Framework and will include Leaflet and D3 libraries.

Based on this input data, a choice model can be set up using a logit model which is common practice in transportation modelling (Bierlaire 2016). The output should show the influence of the above-mentioned attributes (start-time, duration, change-overs, leading and lagging walking distance) on the order of the results from the routing algorithms. An initial evaluation of the relevance of the results is important as this procedure is extremely novel. In the past, research has been conducted on how people are influenced by an order of alternatives but not on how an algorithm influences the order of alternatives. This is where reverse engineering starts, because choice modelling is usually attributed to decisions by individuals rather than algorithms.

REFERENCES

Bierlaire, M., 2016: PythonBiogeme: a short introduction

CITY VIS
WORKSHOP

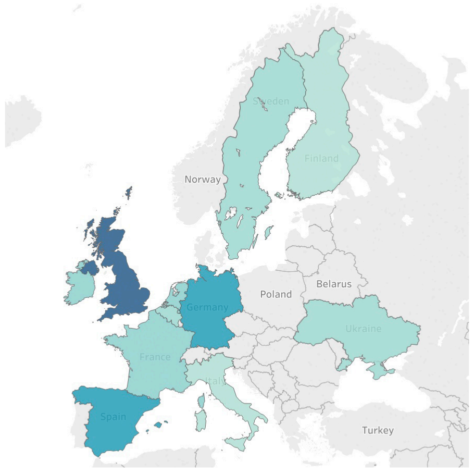


Figure 1: Number of visualisations assessed as per country

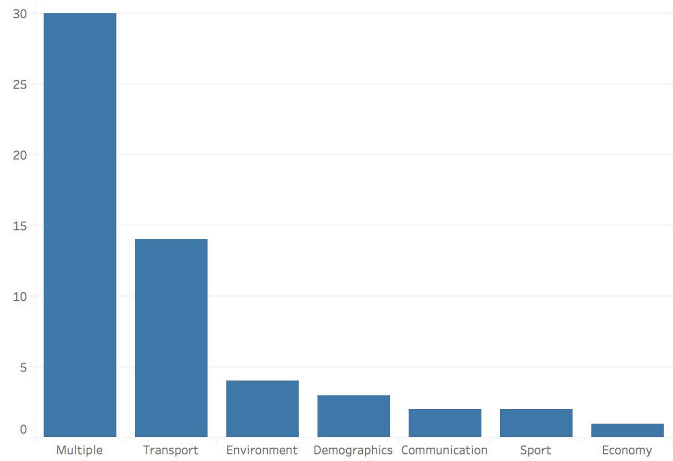


Figure 2: Main topics visualised

Heike Vornhagen[1] (heike.vornhagen@insight-centre.org), Brian Davis[2] (brian.davis@mu.ie),
Manel Zarrouk[1] (manel.zarrouk@insight-centre.org)
[1] Data Science Institute, NUI Galway, [2] NUI Maynooth

SHOW ME THE CITY!
MAKING SENSE OF VISUAL REPRESENTATIONS OF CITIES WITHIN THE EUROPEAN CONTEXT

Keywords: Human-centered computing—Visualization; Applied computing—Computing in government; General and reference—Evaluation

1. RESEARCH PURPOSE

This research aims to establish an overview of current city visualisations and how these depict cities as complex systems. This is informed by a systems thinking approach as we are convinced that there is an urgent need to communicate cities as complex systems to a variety of users.

2. BACKGROUND

Visualisations about all aspects of city life are increasingly used to navigate and understand city life especially in the Smart City context - yet little research has been undertaken as to their impact on users. Another important aspect regarding city visualisations is a need to fully capture cities as complex systems in order to enable better understanding of the impact of decisions across the city. Yet most visualisations in the city context focus only on a part of it, or - as in the case of governance dashboards - enable access to a variety of data about a city without giving users the means to make sense of the city as a complex system.

2.1 CITIES AS COMPLEX SYSTEMS

Cities are good examples of complex systems because the different spheres of city life interact with each other in a complex interchange across city departments. These relationships are not necessarily visible or logical. Furthermore, cities are situated both within a temporal/historical and spatial context, each of which may add another layer of interaction. They contain both ordered / planned elements and unordered, organic elements, areas that can be defined and measured and those that cannot - and perhaps should not [2].
Complex systems are intrinsically difficult to model and very difficult to fully understand. According to [5] humans are conditioned to understand systems through a “narrow, event-oriented and reductionist” focus, which is compounded by a fragmented view of the world around them - and these “mental boundaries” are so ingrained that they are considered natural and right. Factors such as time, hidden relationships and internal dynamics often influence system behaviour [6] - yet, peoples tendency to

oversimplify complexity, and to expect quick-fix solutions [3] may lead to detrimental outcomes.
Visualisations have the potential to explain these complex issues to a variety of users and city-focused visualisations have began to be developed not only by visualisation designers but city authorities themselves.

2.2 VISUALISING CITY SPACES

Visualising city spaces is increasingly seen "as a tool for analysis, exploration and communication" which "has become a driving force in the task of unravelling the complex urban fabrics that form our cities"[1]. However, visual representations of reality are already an interpretation of that reality [4]. Without an understanding by the user of the underlying methodology, visualisations end up contributing to a defined epistemology. This may be less of an issue for specific topic-limited areas, but in the case of complex systems such as cities, this 'reductive/limited knowledge creation will result in skewed pictures of reality. Decisions made based on this skewed picture are bound to deliver not what is really needed but what is perceived to be needed.

3. DESIGN / METHODOLOGY / APPROACH

In order to understand more fully the variety of visualisations regarding cities we carried out a mapping exercise capturing visual representations of city topics in the European context. With the European Commission's launch of the Open Data Strategy and subsequent guidelines, access to city data at EU level has become easier, and, consequently, creating visualisations using this data has increased (visualisations surveyed for this study are predominantly post-2011). We deliberately did not limit ourselves to visualisations that explicitly focused on complexity for a number of reasons:

1. Visualisations focusing on just one topic (e.g. transport) may still contain elements of wider city issues,

2. 'Simple'visualisationsmaybeseenbyusersinacitycontext irrespective of initial aim by the creator,

3. City visualisations may contain useful methods that could be employed within a visualisation of complexity,

4. Thestudyaimstoprovideanoverviewandattemptstogroup visualisations by different criteria, therefore visualisations that do not focus on complexity are still valuable components in the overall picture.

We searched for 'City Visualisations', 'Visualising Cities', 'City Dashboards', 'Urban Data Visualisations' and assessed the visualisations according to the following criteria:

• Type of visualisation (Map, Infographic, Arts, Film, 3D, Dashboard)

– Map included spatial maps of any kind (geographical maps, representative maps eg. public transport network, voronoi maps)

– The term 'Infographics' was used to describe visualisations that combined a visual of any kind / a number of graphs with substantial text (eg a newspaper story)

– The term 'Arts' was allocated to visualisations that look at city space from an artistic perspective and includes for example installations

– Film - a standalone video describing a city related topic

– 3D-city visualisations that used only 3D

– Dashboards - amalgamation of data about a city with some visual exploration possible

• Location - Global/City/Country/Region (as appropriate)

• If the visualisation was exploratory or explanatory (y/n)
Any visualisation that allowed user interaction was deemed to be exploratory while those visualisations that included textual explanations to support sense making of the topic were classified as explanatory

• If the visualisation offered users to download data used (y/n)

• Topics covered (transport,housing,environment,etc.)

• Complexity (y/n)

A visualisation was deemed to describe complexity if it enabled users to compare somewhat unrelated topics (e.g crime and housing).
Visualisations about cities outside Europe were generally excluded though projects that had both European and non-European aspects were included. For example, a visualisation may depict public transport usage in Berlin, London and New York. This was included and tagged as 'Global'.

4. ORIGINALITY / VALUE

In order to develop a visualisation model that better captures cities as complex systems, we first need to develop a baseline of existing city visualisations and how these depict complexity. This builds on a recent study which reviewed city visualisations within computer science literature [1] and extends / develops assessment criteria to judge the potential for an exploration of complexity (e.g. interactivity, context, cross referencing). [1] reviewed 32 computer science papers and assessed the therein contained city visualisations according to topic and type. They found that traffic is the topic most often covered, while maps are the most often used type. However, they did not in any way assess the visualisations further in terms of user interaction or complexity. To the best of our knowledge, no study of this kind has been attempted before.

5. LIMITATIONS

The study focuses on the European context only and is limited by the use of British English (partly in order to exclude US based city visualisations). Other European languages are currently being researched in order to capture visualisations in non-English speaking countries. Furthermore, while the EU guidelines apply only to EU member states, we envisage a push for similar initiatives in other European countries and hope to extend the study respectively.

6. PRACTICAL IMPLICATIONS

So far we have reviewed 63 visualisations mainly of Western European cities[2]. Of the visualisations depicting single topics, transport is the most common. As such, most visualisations assessed so far depict more than one topic e.g transport & sentiment.

However no visualisation was found that enabled users to make sense of a city in its totality. While dashboards do offer access to data to a range of topics, comparing them across different sectors remains difficult. Furthermore, as outlined above, the data made available does not necessarily depict all aspect of a city.

7. IMPACT

The results from our search so far indicate that a visualisation de- picting a city in its complexity does not currently exist. Most visu- alisations about cities focus on specific topic areas and while some combine 'unrelated' areas these rarely go beyond a handful of topics. Dashboards, while collating data about a myriad of city related areas, also rarely allow users to query 'unrelated topics - e.g. it is extremely difficult to find information on how traffic (transport) may impact mental health (health). Considering the impact of policy decisions at local level it is important that a way is found to successfully capture a city as a complex system.

ACKNOWLEDGMENTS

This work has emanated from research supported in part by a re- search grant from Science Foundation Ireland (SFI) under Grant Number SFI/12/RC/2289 and in part by the SSIX Horizon 2020 project (grant agreement No 645425).

FOOTNOTES

[1] "CityVis.io", http://cityvis.io/about.html accessed Nov 10th 2017

[2] See for example http://422south.com/work/invisible-netherlands, http://map.glasgow.gov.uk, http://galway-dashboard.ie http://imcg.se/en/project/inner-citys-energy, http://metropolitain.io or http://mappingforchange.francastillo.net

REFERENCES

[1] A. Eberhardt and M. S. Silveira. Show me the data! a systematic mapping on open government data visualization. In Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age, dgo '18, pp. 33:1–33:10. ACM, New York, NY, USA, 2018. doi: 10.1145/3209281.3209337

[2] R. Kitchin. The real-time city? big data and smart urbanism. GeoJournal, 79(1):1–14, 2014. doi: 10.1007/s10708-013-9516-8

[3] D. A. Norman and P. J. Stappers. Design x: Complex sociotechnical systems. She Ji: The Journal of Design, Economics, and Innovation, 1(2):83–106, 2015. doi: 10.1016/j.sheji.2016.01.002

[4] B. Rieder and T. Roehle. Understanding Digital Humanities, chap. Digital Methods: Five Challenges, pp. 67–84. Palgrave Macmillan, London, 2012. doi: 10.1057/9780230371934 4

[5] J. D. Sterman. All models are wrong: reflections on becoming a systems scientist. Systems Dynamics Review, 18(4):501–531, 2002.

[6] J. D. Sterman. System dynamics: Systems thinking and modeling for a complex world. pp. 1–29, 2002. doi: 10.1287/mnsc.35.3.321