N. Gu, S. Watanabe, H. Erhan, M. Hank Haeusler, W. Huang, R. Sosa (eds.), *Rethinking Comprehensive Design: Speculative Counterculture, Proceedings of the 19th International Conference on Computer-Aided Architectural Design Research in Asia CAADRIA 2014,* 45–54.

INTEGRATING USER AND USAGE INFORMATION IN A DESIGN ENVIRONMENT

ANDRÉ CHASZAR[1] and BIGE TUNÇER[2]
[1,2] *Singapore University of Technology and Design, Singapore*
{andre_chaszar, bige_tuncer}@sutd.edu.sg

Abstract. We describe research exploring and demonstrating the use of large-scale data gathering and processing to inform the activities of architectural and urban designers. We apply this research to public spaces in urban housing estates. The aim is to understand the current use patterns and usability of these spaces, and to adaptively redesign them according to the insights gained from these findings. Another aim of the research is to obtain scientific knowledge regarding the production and use of user-data-based design support systems which promote and enhance the capability of (digital) design aids, such as building- and urban-scale models, to act as 'learning devices' giving designers better insights to the nature of the design situations they are asked to address, as well as insights on design space definition and exploration. We adopt a multi-modal data collection strategy, consisting of participatory workshops for residents and users, person-based crowdsourcing, location-based crowd sensing, and statistical demographics data.

Keywords. integrated design environment; multi-modal data collection; data visualization; data analysis; public space design.

1. Introduction

The design of urban spaces, particularly open public ones, is a complex subject approached in research to date most often from traffic management perspectives such as pedestrian flow modelling. The increasing availability of various kinds of data on actual use of public spaces is making more feasible a data-based approach which can perhaps replace or at least supplement theory-based approaches. However, hurdles remain in both knowledge and application regarding integration of space use data into design activity. In this paper we present an account of steps we have taken in defining and address-

ing important issues for research on this subject, and we also present some preliminary results of our research activities to date.

At the present stage of work we are assuming the availability of data addressing various dimensions of public space use, although subsequent work will also address issues regarding the particulars of what kinds of data are available, how they are collected, what additional kinds of data would be desirable for informing design activities, and how these latter could be collected and made available. Given an adequate supply of data, challenges of particular interest include: first, the aggregation and disaggregation of data, for example through mapping data points to (physical) spatial locations, and second, the representation of data in ways which maximally support and minimally hinder the designers' workflows and thought processes. Techniques to be examined and compared include location-based vs. person-based data gathering, mapping of data to (predetermined) objects vs. to (emergent) regions, data visualization via a design-oriented vs. primarily statistical- or other data mining-oriented software platform, and various HCI techniques to facilitate the designers' interactive access, analysis, evaluation and response to and of the gathered data.

A significant sub-component of the project's person-based data-gathering aspect comprises development and deployment of a mobile 'app' for voluntary solicitation of participating persons' descriptions of and reactions to the physical environments under design study while they are in them. This in turn engages a sociological dimension of the research, wherein non-quantifiable (or only roughly quantifiable) factors are considered as relevant and necessary data in the evaluation of design performance and quality. Information modelling constitutes another important component of the research, in its bridging between data collection, processing, (re-)presentation and communication, and in its conditioning of the mappings which may be needed between these various modes of data use and knowledge generation. The research also includes exploration of various degrees and levels of interactivity between designers and the gathered data, ranging from, for example, the provision of user-adjustable data filters for displaying pre-programmed analysis results, to provision of alternative data analysis methods from which the designer/user may choose. These in turn impact upon the power, usability, required expertise, learning potential, etc. of the data tools. Outcomes of these comparisons are to be tested in various ways including user-testing of tool implementations in the course of development, as well as eventual application of design knowledge gained via production of a small intervention within the design study area and subsequent analysis of occupant/users' reactions to and use of the new facility.

In the following sections, we describe in more detail the research goals,

background, research methodology and preliminary results, concluding with a discussion of these and indications for further work.

2. Research goals and background

Our research on this topic is motivated by the availability on one hand of ever increasing amounts and types of data regarding usage of the built environment, and on the other hand an expressed desire and need for designers of the built environment to use these data for improving their insights to usage and also to improve the quality of their designs and the resulting environments. Urban open space in particular is an area of great interest for various reasons, ranging from commerce and cultural exchange through security and community bonding, and morphological studies open one angle of approach to this subject (Moudon, 1997) by positing the existence of generalisable relationships between the physical forms and other qualities of spaces and the uses of spaces, seeking or hypothesising specific examples of such relationships, then in some cases investigating the extents to which the discovered or assumed relationships hold true. Efforts to model (e.g. via simulation) the usage of public spaces tend to focus on analyses of movement, such as Hillier and Hanson with 'space syntax' (1984). However, the quality of urban space depends also on recognition of space as a place of occupation, not only of transit. Such phenomena have also been studied (e.g. Whyte, 1980) and are closer to the concerns of architects and urban designers than of traffic engineers.

Technical issues concerning integration of 'big data' into software environments such as BIM and GIS have been partially addressed by some researchers and practitioners as detailed in Section 3. This set of concerns involves not only data transmission (to the machinery hosting the software) but also and more crucially data representation (to the users of the software) and the corresponding issues of HCI (human-computer interaction) and cognitive engineering (Feltovich et al, 2004).

Aspects of sociological research methodology include questions of more structured or less structured approaches to gathering data pertaining to public space use, relying primarily or entirely on collection of data within existing categories of knowledge versus taking a more grounded, data-driven approach where factors of importance emerge from the data themselves (Glaser and Strauss, 1967). These issues are important in relation to objective, quantifiable as well as subjective and other 'soft' data which include much relevant information beyond mere physical quantities.

Our work thus aims to tap into the large data stores on public space usage being gathered via various sensors, augment these data where appropriate to

complement and corroborate them with information not readily available by passive sensing, and analyse and present these in ways useful to designers so as to enable their formation and testing of hypotheses regarding use of public space, confirming their existing knowledge or revealing new insights.

3. Methodology and approach

The research follows an inductive approach, and brings together technological research, design research, and sociological research. The overall research methodology that is adopted is grounded theory, which aims to develop theory that is grounded in data systematically gathered and analysed (Martin and Turner, 1986). Additionally, various methods apply to various aspects of the research.

Data collection will occur through a multi-modal approach, mainly space-centric and people-centric crowdsensing, telco data, and participatory workshops. One of the concerns in data collection is ensuring that the data samples are representative of the diverse demographics of the population in the studied area. Apps that collect data from users will require the use of smartphones, and therefore the data collected has the risk of being biased towards a certain demographic group, especially excluding the elderly. Therefore supplying certain demographic slices with smartphones will be part of our methodology. Literature is available demonstrating donated mobile phones and free apps to the elderly to use and has been satisfactorily effective (reference to be added after the blind review phase). This indicates that it is possible for elderly to use smart phones in their daily life.

The crowdsourcing apps on the smartphones will serve two purposes. One will be for the collection of GPS-based location information. The other purpose is for the users to provide their opinions and feelings about the specific public spaces and their use in a playful and fun manner. We additionally intend to entice all users to download and use the apps intended for data collection for this project. A few issues are of great importance here: usability of the apps, advantage for users, and feedback to users. The apps must be easy to use, and must work in several modes. They may run in the background for GPS data purposes. If users wish to engage in subjective feedback activity about public spaces, apps must provide easy to use and appealing interfaces. Finally, users must be kept up to date about how their participation is important and why the data will be useful for the adaptive design of their public spaces.

The space-centric sensing will provide us with data about the usage of public spaces and movement count of people on pathways. This requires the installation, configuration and maintenance of sensors and sensor network.

The sensors need to be programmed so that the desired sampling is achieved in a power efficient manner, and a wide area of coverage is obtained.

The telco data enables us to track a device in relatively coarse location resolution for a time period, e.g., one day. Our analysis of the data will focus on extracting and interpreting the insights at the population level, which eliminates the personal privacy issue. This method provides reliable and consistent data over time. The participatory workshops will include all slices of the demographic structure. These workshops will be conducted with several stakeholders, among others, users of public spaces, designers, and policy makers. When collecting data from residential areas, users are mainly residents. In other, commercial or transit areas, users are visitors and people who work near these areas, as well as those stationed there and/or responsible for their upkeep, such as maintenance and security staff, facilities managers and the like. The multi-modal data collection approach is an excellent method to collect data about users' evaluation of the quality of spaces. Merging and analyzing all this data together will allow designers to form an informed opinion about what issues to take into account when designing public spaces.

3.1. RESEARCH COMPONENTS

The research is conceived as comprising four main components:

1) *Research foundation*, including the sociological framework and spatial analysis and mapping subcomponents;
2) *Testbed deployment and data collection*, with subcomponents such as statistical demographic data, location- and person-based sensing and observation, telecommunications data, and participatory workshops;
3) *Data processing and analysis*, including data preprocessing and integration, activity recognition and usage pattern analysis, and mobility analysis; and
4) *BEM, prototyping and evaluation* with a Built Environment Modelling (BEM) platform, adaptive public space design prototypes, and design evaluation subcomponents.

These compose a roughly sequential cycle of workstages, with results of the fourth modelling, prototyping and evaluation stage being fed back into the initial research foundation, although with some other subcycles in the process also anticipated (Figure 1).

3.2. RELATED WORK

Work on integration of data – especially large, external data sets (such as sensor data) – into design environments has been proposed and in some few cases implemented in practice as well as in academia. In the realm of aca-

demic research, integration of data (not necessarily "big data") into design-analysis processes occurs both as an incidental component of other research aims which simply have need of the data, and also as a topic of study in its own right. The design-analysis processes may be situated in various platforms, including diverse CAD and CAE applications, as well as BIM. Akinci et al (2007) describe problems arising in integration of data from GIS sources into CAD platforms, and propose semantic web-based services as a way forward. Hijazi et al (2011) take the position that integrating capabilities of CAD/BIM and GIS systems is still not feasible, and so linking them should be approached at the data level, as also indicated by Wu and Hsieh (2008). More recently, van Berlo and de Laat (2011) address integration of GIS with BIM via the CityGMLGeoBIM extension, in terms of mapping comparable objects but without specific use cases. These discussions touch also on the issue of light- vs. full data exchange standards, e.g. for IFC, and whether sensor data is best integrated to design environments via existing data formats or more directly. Lapierre and Cote (2007) discuss urban data management on a testbed for CAD/BIM/GIS integration via open web services. Dealing more specifically with sensor-derived data, Ganti et al (2011) identify existing capabilities and challenges for mobile "crowdsensing". Meanwhile Phua et al (2011) describe methods for using simple sensors to recognise patterns of activity.

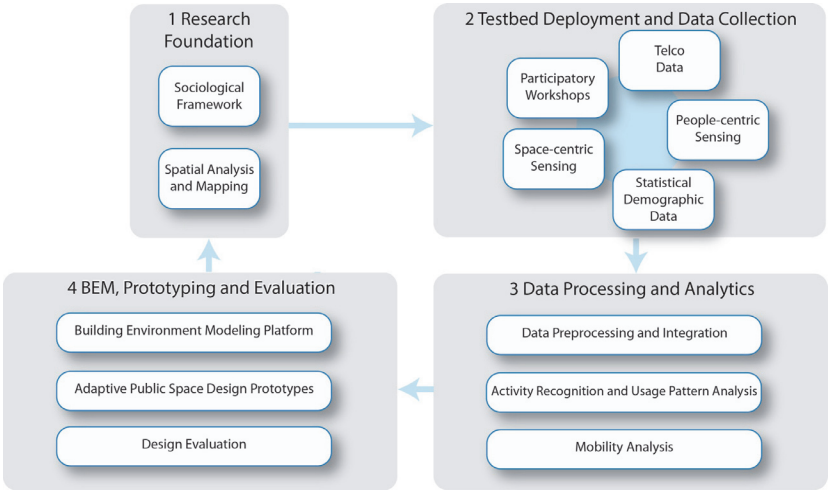

Figure 1. The overall approach and methodology of the research.

Functioning implementations of cross-platform data exchange including various BIM environments is offered commercially by vendors such as G-Team, with capabilities to handle a variety of 2D and 3D geometric as well

as some non-geometric data formats. Field sensing of construction work to assess progress, accuracy and precision of installed items is also an area of ongoing application development. Integration of field data and other sources within a BIM-based environment has been implemented and used in practice via the Dunn Dashboard, a cloud-based collaboration platform capable of handling digital models and documents as well as photographs and video (Barista, 2013).

Regarding data visualisation and analysis, the field of Visual Analytics comprises various methods for processing and presenting large data sets, of which Keim et al (2010) give a thorough overview. Tufte and Graves-Morris (1983) meanwhile provide guidelines and case studies for informative versus unclear visualisation design techniques and strategies, applicable to both digitally and non-digitally derived and presented data.

The sociological dimension of our topic (the design of urban open public spaces using data to inform design decisions) is addressed by some researchers, though much of the implementation in practice remains relatively undocumented in literature. One outstanding counterexample is the body of work by Whyte et al. (1980) which studied persons' individual and interactive behavior in a variety of urban publilc space settings, drawing conclusions about features and/or characteristics of such spaces which promote, hinder, or otherwise seem to influence use of the spaces. Ideas about the relationship of spatial design to human behavior in urban settings also motivate the work of Hillier and Hanson (1984), for example in the Space Syntax method of analysing movement in network representations of streets and other open spaces. Jacobs' (1961) observations of urban street life also shed light on relationships between designed/constructed forms, regulations and use of space. Lynch (1960) also derives conclusions regarding spatial qualities and people's comprehension, navigation and other use of urban fabric. It is worth noting, however, that the data in the work of Whyte, Lynch and Jacobs is obtained through field observations, not sensing. Integration of observational data, sensor data and contextual data (such as social media feeds) is a challenging problem relating to ontologies and folksonomies, with semantic web-based methods again finding relevance (Angeletou et al 2007).

Liv(e)ability is commonly invoked as a goal for improvements in the quality of public space and as a motivation for technologies to support better design therof. The term’s current use includes some conflation with the term sustainability, e.g in van Dorst (2012). However, its emergence in the 1970s and relation to New Urbanism as a movement independent from environmentalism – with quality of construction, and quality of life issues among its foremost concerns – can be at odds with the latter, for example by including 'greening', but not necessarily consistently with aims of sustainability

(Bocquet, 2013). Thus some clarification of the actual end goals is needed to better design the technologies meant to help achieve them.

Within this framework of relevant precedents our research project aims to increase the availability and comprehensibility of actual (sensed and volunteered) data on space usage to designers, enabling in some cases confirmation of existing knowledge and hypotheses/beliefs about the effect of spatial configuration on usage, or in some cases discovery of new patterns.

4. Preliminary results

While much of the work we have described so far lies in the future, some experimentation is already in progress and some results are already available (with more anticipated by the time of the conference presentation). Among the more prominent aspects of the research is the user interface, both in terms of visibility and of significance, as mentioned above, as this element has a crucial influence on the quality of insights derivable from the data.

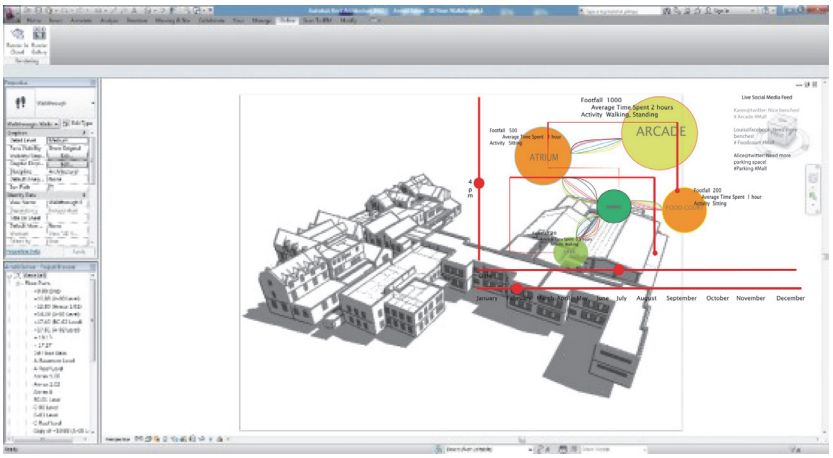

Figure 2. Mock-up of one possible envisioned visualisation of public open space user and usage information within an architectural design environment.

An example data visualisation figure (Figure 2) shows how one such interface might look, superimposing various data such as footfall, average time spent and predominant activity types at particular time intervals (e.g. day of week, hour of day, etc.) onto a geometric representation of the spaces being examined. The data include both sensed and derived items, with raw data from sensors having been aggregated and partially interpreted according to pre-determined analysis procedures (not designer-specified ones, though these could also be accessed through menus, toolkits, etc. if available). The

mapping of data here is onto predetermined spaces or regions, such as 'arcade', 'atrium', 'food-court', etc. General movement and usage data thus aggregated is supplemented with a (live or historical) feed from sources such as social media, news, and field observations which can provide additional context and/or detail not readily captured in the numerical and activity-type/-location data.

In addition to the design-view type of visualisation shown here, it may be necessary to develop intermediate views which deal more directly with the raw (or more nearly raw) sensor data, in order to enable revisions of the aggregation and analysis, for example redefinition of the boundaries between labelled 'activity zones', where these are not clearly delimited.

5. Conclusion

We have described above our motivations and approach for a study of how data on use of public space can be integrated into a design environment in order to enable designers to gain insights and thus make better design decisions about new public spaces or modifications of existing ones. We have also indicated some of the related work already carried out in previous research. Lastly we have described some preliminary results which we have obtained using the information and other resources available to date. Our plans for continuing work on this topic include more detailed definition of the anticipated data collection needs and available capabilities as well as further work on design and technical issues relating to analysis and visualisation of the data by and for users.

We see the significance of this research in the context of current activities concerning Big Data and Smart Cities. In academia, for example, research by the MIT Senseable Cities group and the SEC Future Cities Laboratory address aspects of data-driven gathering of insights for design and decision making, while in practice some cities have already taken various steps for sensing data in relation to issues such as transportation, energy use and crime. However, these other efforts do not sufficiently address the specific issue of urban public space quality, nor the challenges of data collection and processing particular to it. Furthermore, an open question remains regarding how to analyse and evaluate cities' sensed big data so as to proceed from description to action. This requires among other things the establishment of appropriate metrics to indicate when intervention is needed and possibly also what options for intervention are available, as well as the testing of what-if scenarios. Our research thus can contribute to this specific domain of inquiry (the quality of public space), and also to investigating and developing knowledge and methods necessary for the more general problem of how to

draw actionable conclusions from the large quantities and diverse types of data collected, to make the data more valuable and support the expenditure of effort and other resources necessary for obtaining them.

References

Akinci, B., Karimi, H., Pradhan, A., Wu, C. C., Fichtl, G.: 2008, CAD and GIS interoperability through semantic web services. *ITcon* 13, 39–55.

Angeletou, S., Sabou, M., Specia, L. and Motta, E.: 2007, Bridging the gap between folksonomies and the semantic web: an experience report, *The 4th European Semantic Web Conference 2007 (ESWC 2007)*, 3-7 Jun 2007, Innsbruck, Austria.

Barista, D.: 2013, "'BIM for all' platform pays off for comtractor". Available from: Building Design + Construction <http://www.bcdnetwork.com/'bim-all'-platform-pays-contractor /> (accessed 13 December 2013).

Berlo, L.v., and Laat, R.d.: 2010, Integration of BIM and GIS: the development of the CityGMLGeoBIM extension. In T. H. Kolbe, G. König, & C. Nagel (Eds.), *5th International Conference on 3D GeoInformation. XXXVIII-4*. Berlin: ISPRS.

Bocquet, D.: 2013, Liveability vs Sustainability? Ambiguities in Present Debates about the Future of Our Cities, public lecture at LKY Centre for Innovative Cities, 4 Dec.

Dorst, M.v.: 2012, Liveability. *Sustainable Urban Environments*, Springer, 223-241.

Feltovich, P. J., Hoffman, R. R., Woods, D., and Roesler A.: 2004, 'Keeping it too simple: how the reductive tendency affects cognitive engineering', IEEE Intelligent Systems, May/June, , 90-94.

Ganti, R. K., Ye, F. and Lei, H.: 2011, "Mobile Crowdsensing: Current State and Future Challenges, " *IEEE Communications Magazine*, **49**(11), 32-39.

Glaser, B. G. and Strauss, A. L.: 1967, *The Discovery of Grounded Theory: Strategies for Qualitative Research*.

Hijazi, I., Ehlers, M., Zlatanova, S., Becker, T., and v.Berlo, L.: 2011, Initial investigations for modeling interior utilities with #D geo context: tranforming IFC-interior utility to CityGML/UtilityNetworkADE, Advances in 3D Geo-Information Sciences, 95-113.

Hillier, B. and Hanson, J.: 1984, *The Social Logic of Space*.

Jacobs, J.: 1961, *The Death and Life of Great American Cities*, Random House, New York.

Keim, D., Kohlhammer, J., Ellis, G., and Mansmann, F. (eds.): 2010, *Mastering the Information Age: Solving Problems with Visual Analytics*, Eurographics, Goslar.

Lapierre, A. and Cote, P.: 2007, Using open web services for urban data management: a testbed resulting from an OGC initiative for offering standard CAD/BIM/GIS services, *Annual Symposium of the Urban Data Management Society*, 381-393.

Lynch, K.: 1960, *The Image of the City*, MIT Press, Cambridge.

Martin, P.Y. and Turner, B.A. (1986). Grounded theory and organizational research. *Journal of Applied Behavioral Science* 22(2), 141-157.

Moudon, A.V.: 1997, Urban morphology as an emerging interdisciplinary field, *Urban Morphology*, 1, 3-10.

Phua, C., Sim, K., and Biswas J.: 2011, Multiple People Activity Recognition Using Simple Sensors, *Proceedings of PECCS11*, 224-231.

Stouffs, R.: 2001, Visualizing information structures and its impact on project teams: an information architecture for the virtual AEC company. *Building Research & Information*, **29**(3), 218-232.

Tufte, E. R. and Graves-Morris, P. R.: 1983, *The Visual Display of Quantitative Information*.

Whyte, W. H.: 1980, *The Social Life of Small Urban Spaces*, Conservation Foundation.

Wu, I. and Hsieh, S.: 2007, Transformation from IFC data model to GML data model: Methodology and tool development, *Journal of the Chinese Institute of Engineers*, **30**(6), 1085-1090.