

Matthias Scheutz and
Tufts University
161 College Avenue
Medford, MA 02155, USA

Thomas Mayer
Independent Architect
Koppstrasse 103
1160 Wien, Austria

Combining Agent-Based Modeling with Big Data Methods to Support Architectural and Urban Design

10/21/2015

Abstract

Big Data analytics are increasingly used to discover potentially interesting patterns in large data sets. In this chapter, we discuss the potential of combining Big Data methods with those of agent-based simulations to support architectural and urban designs, for agent-based models allow for the generation of novel datasets to study hypothetical situations and thus designs. Specifically, we present two conceptual studies that investigate the utility of agent-based models in conjunction with Big Data analytics in the context of multi-level pedestrian areas and current office designs, respectively. The analyses of the case studies suggest that it will be worthwhile, both for urban designers and architects, to pursue a combined agent-based simulation Big Data analytics approach.

1. Introduction

The notion of 'Big Data' is becoming increasingly popular these days and statistical data mining or machine learning techniques applied to 'Big Data' sets have already been used in various application domains, from research settings in physics and genomics (e.g., Costa 2012), to industry and government (Cull 2013). While there is no precise formal definition of what is typically intended with the term, paraphrasing Kenneth Cukier and Viktor Mayer-Schönberger (Mayer-Schönberger & Cukier 2013), we take 'Big Data' to refer to the kinds of discoveries of principles and relationships that are possible with sufficiently large data sets that would not have been possible with smaller data sets.

Big Data can thus be used in many areas for the discovery of new, often unexpected qualitative and quantitative relationships as in the case that made the news a few years ago when the US Target Corporation used 'Big Data methods' to predict the pregnancy of costumers based on the costumers' shopping behaviors (Duhigg, 2012). Note that by 'Big Data methods' we do not intend to refer to the various computational methods and challenges for organizing, storing, and retrieving data from large distributed databases, but

rather to the various statistical and algorithmic methods for analyzing large data, detecting patterns and regularities, and extracting higher-level relations, rules, and principles from it (cp. to Gandomi & Haider 2015).

An exciting new application area of Big Data has been evolving around the theme of 'urban computing', with a focus on developing data mining and data fusion methods for large available data sets from municipalities and other government sources to determine principles and relationships in the context of urban areas (e.g., effects of traffic patterns on air quality distributions, the functions of different urban regions, patterns of consumption and the effects on energy distribution, or traffic patterns and hold-ups). One goal of *urban computing* is to provide solid data analyses that might help urban designers make better fact-based design decisions that would not be possible without the new insights gained by combining analyses from different data sources. Whether and to what extent Big Data analyses can actually help urban designers, and possibly also architects, in improving their designs is, however, still an open question. For one, it is not clear whether and how results from the performed data analyses (e.g., such as the result from analyzing the relationships between air pollution and traffic patterns) could be translated into workable solutions (e.g., to improve air quality or traffic flow). The difficulty arises from the fact that urban designers have to take multiple constraints into account, some of which may not be addressed by the results from the performed Big Data analyses (e.g., data from air pollution based on traffic pattern is not the only factor used in deciding whether to develop a region or place a new airport).

Consequently, some have recently called for an encompassing new *Urbanization Science* that will be able to systematically investigate, inter alia, the question about the utility of Big Data in urban designs (e.g., see Solecki et al. 2013). The importance and urgency of this question is also evidenced in the various new initiatives and research centers that have been formed in the very recent past specifically to investigate the use of Big Data for urban planning and development. For example, in 2012, the *Urban Center for Computation and Data* was formed jointly by the University of Chicago and Argonne National Laboratory in the US with the goal to apply computational and data-driven methods for urban planning and design (see CCD, 2015). Other related centers that were recently created with the goals of putting available cloud-based services and Big Data to use in urban planning and design are the *Center for Urban Science and Progress* (see CUSP, NYU, 2015), the *Urban Systems Collaborative* (see Urban Systems Collaborative, 2015), and Microsoft Research's *Urban Computing Group* (see MS Urban Computing, 2015).

We believe that it will be important to combine Big Data methods with more traditional simulation methods, in particular, *agent-based* modeling. This will allow researchers and practitioners to utilize the strengths of both approaches in order to explore design spaces and options in an unprecedented manner. Of course, this will require both solid

computational frameworks for integrating large data sets and analyses with simulation frameworks, as well as appropriate user interfaces and data visualization tools to allow non-experts to work with complex simulations effectively, but fortunately these types of computational and infrastructure challenges are already being addressed by the computer science community.

The goal of this chapter is thus to first analyze whether and to what extent Big Data methods can be combined with more traditional simulation-based methods, particularly, those involving *agent-based models*, and second whether consequently such a combination could improve the designs of urban environments and living spaces, from designs by urban planners to designs by architects. To our knowledge, there have been no attempts to apply Big Data methods to large data sets generated by simulation runs of agent-based models, let alone in urban computing, which then use the results to guide model refinement and further model development. By providing two case studies that highlight the potential but also limitations of what we believe is a promising combination of mathematical and computational methods we hope to set the stage for modelers and designers to embark on future projects that explore the synergistic potential of Big Data methods and agent-based simulations.

2. Combining Big Data and Agent-Based Modeling for Urban and Architectural Design

2.1. Agent-based Modeling

Agent-based modeling (ABM) is an approach to the study of complex systems where the laws guiding the overall system behavior are unknown, while the laws guiding the behavior of parts of the system are known. The modeler then defines those parts together with their behaviors and runs simulations of their interactions to study how those parts affect the behavior of the whole system.

Agent-based modeling has been successfully employed for quite some time in the behavioral and social sciences (e.g., Grimm et al. 2006, p. ?). In our own past work, for example, we have developed a distributed agent-based simulation environment *SimWorld* to explore a variety of models of artificial life scenarios, evolutionary trajectories, social simulations, swarm-based simulations and individual-based biological models of group behavior, and various others. (This simulation environment was implemented in our ADE-GRID computational middleware which supports systematic explorations of large model parameter spaces that yield insights in actual and counter-factual models (e.g., see Scheutz & Harris 2011)). Other examples of agent-based modeling relevant to urban systems are models of walking agents and groups of agents, which are often based on *force dynamics* to guide the individual agents' behaviors (e.g., see Torrens & McDaniel 2013, or Ali et al. 2013).

The ability of agent-based models to discover designs that might seem counter-intuitive at first is a feature of agent-based models that is of particular relevance to urban

systems, especially because such discovered designs can often be empirically verified (cp. to the model-based prediction of placing obstacles near exits to allow crowds to evacuate an area more quickly in an emergency situation, Yanagisawa et al. 2009).

Agent-based modeling has a main advantage over other forms of modeling such as mathematical modeling using difference or differential equations to capture law-like regularities among state variables or more abstract non-mathematical forms such as 'verbal models' which describe relationships informally as rules or principles in natural language: complex behavior can be viewed as an 'emergent phenomenon' resulting from interactions among many individual agents. And this type of emergent behavior is typically not predictable from initial model conditions. In fact, we can mathematically prove that even the simplest such simulations do not allow for predictions of future states (e.g., the 'game of life,' Grim et al. 1998); instead, we have to run model simulation in order to determine whether a particular model state can be reached from a given initial state.

Rather than requiring interaction principles at the group level, which are often unknown, agent-based modelers specify rules for the behavior of individual agents, which are typically easier to come by (e.g., how a pedestrian moves along a sidewalk). Once the behaviors of the agents are fixed, simulations of the model can be run in a variety of conditions to explore the dynamics of the interactions among the agents and the evolution of the model states over time (e.g., how crowds can spontaneously form in some locations based on the number of agents, their moving directions, or the layout of the sidewalks). The resultant complex behaviors can then often be captured in terms of higher-level group variables (e.g., clusters of agents), and laws governing the changes of those variables over time can be derived from analyses of the simulation results (e.g., clusters appear, move, and disappear based on time of day). Agent-based models can also capture extreme heterogeneity among agents and allow for very flexible modeling of spatial characteristics of the environment (such as the distribution of agents), both features that are important for urban design. Moreover, agent-based models can be easily extended to multiple levels of organization and control, which are also critical aspects for modeling complex urban systems (e.g., individual human agents, buildings and their residents, districts, or urban regions). However, agent-based models can also generate large amounts of data, which can be difficult to analyze and understand. Hence the question arises whether agent-based models could be combined with Big Data methods in a way that helps address this problem.

2.2 Big-data Methods

The main strength of Big Data methods is that they can generate new results (concepts, relations, correlations, etc.) from large data sets that would not have been considered and could not have been discovered otherwise. The techniques employed to extract information

from Big Data sources vary depending on the type of data. For example, for text-based data sources techniques include methods for recognizing named entities and relations among those entities, while for sound-based data sets auditory analytic methods (such as speech recognizers) are employed to extract information. Similarly, for video-based data vision processing algorithms are employed to detect and index individuals and objects of interest. Many of these methods are combined in mixed data sets such as social media to determine relationships among the detected individuals (e.g., social influence analyses are aimed at determining an individual's influence within a social network). Moreover, predictive analysis methods (such as interpolation or correlation methods) are employed to generate predictions and trends from historical data. Critically, the classical notion of statistical significance is abandoned in the context of Big Data in favor of model fitting and machine learning methods that directly uncover structure in the data (for more details, see Gandomi & Haider 2015).

2.3 Combining ABM with Big-data Methods

Big Data methods could serve three important purposes in conjunction with agent-based models: (1) they can be used (as data analytical tool) to discover new relationships and concepts relating agents and agent groups in agent-based models that were not anticipated by the model developer, (2) given available large data sets (e.g., data produced by other models, data collected from the modeled domain, or data fused from multiple data sources), they can be used to extract parameters for agent-based models, and (3) they can be used to validate simulation outcomes comparing them to existing available empirical data.

The first case is not only important for analyzing the simulation outcomes in ways that will reveal important relationships that are non-obvious, but also for turning newly discovered qualitative relationships into rules specifying agent behaviors in revised agent-based models, which can then generate new models to explore.

The second case of combining ABM with big-data methods is particularly important for models with a large number of free parameters, e.g., epidemic models of disease spreading. In such models, it is important to fix as many parameters as possible to reduce the overall parameter space that needs to be explored via model simulations. Especially for urban systems, many large, diverse data sets are publicly available from municipalities and can thus be mined using Big Data methods to obtain parameters and validation data for agent-based models.

The third case of combining ABM with big-data methods is important for validating that a model is able to reproduce a known phenomenon at a sufficient level of detail so that model predictions (i.e., model simulations exploring possible future state trajectories of the system under investigation) can be trusted (e.g., how urban areas will grow over the next few decades). Being able to verify agent-based simulation results against Big Data results

enables agent-based models to validate predictions about the target domain that can be used for designs and policy decisions.

The potential use of Big Data methods in conjunction with agent-based simulation models calls for a more detailed investigation of the utility of this combination for supporting urban as well as architectural design, which we will initiate with two conceptual case studies: the optimization of inner-city pedestrian areas in multi-level city centers (which falls under the purview of urban planners) and the optimization of current office designs (which falls under the purview of architects).

3. Conceptual Case Study 1: Optimizing inner-city pedestrian areas in multi-level city centers using Big Data methods

Since the midst of the 20th century secondary and supplementary pedestrian networks have been developed in downtown areas following the 'American Central Business District' model (i.e., a city's business center characterized by high urban density, a concentration of retail, high-rise office buildings, and high-capacity public transport). Cities like New York, Montreal, Toronto and Calgary (the cities this case study is based on) shared the common goal of mastering the lack of capacity of their sidewalks and of unclogging their streets by means of traffic separation. After the simple widening of the sidewalks (and crosswalks) as well as their semi-public extension inside the buildings' perimeter (plazas and atriums), additional pedestrian networks were established by connecting a series of city blocks with either covered bridges across the streets or tunnels underneath them. With these secondary pedestrian networks above or below ground, it became eventually possible to navigate downtown areas without ever setting foot on a sidewalk again (Cui et al. 2013, pp. 151-160, Cui et al. 2011, pp. 1-3).

After a piecemeal and quasi-organic growth of these mostly privately owned networks (in the first half of the 20th century) and the integration of public rapid transit systems (New York's Subway, Montreal's Métro, Toronto's Underground, and Calgary's C-Train), city governments started to establish master plans to guide and direct the secondary networks' future growth. Following the cities' specific nature and history, either tunnels (Montreal, Toronto), plazas (NYC), or covered bridges (Calgary) were favored. So-called bonus systems helped implementing these networks, encouraging private investors to connect their buildings to the network by both allowing more density, extra floors, the use of public ground and even subsidizing the linkages' construction (Department of City Planning NYC 2014, El-Geneidy et al 2011, p. 3, Moore 2013, pp. 2-7, Lucarelli 2014, URA Singapore 2014). After a period of vivid criticism in the 1980ies, focusing on the secondary networks' predominantly commercial character (the shop-lined tunnels and bridges were seen as just another facet of 'the Malling of America' disrupting grown city centers by privatizing public space and sucking street-life

inside their private domain (Crawford 1992, p. 24, Body 1992 p. 124)), the secondary networks' vital role as a pacemaker for the central business districts was recognized (City of Toronto CPD 2012, pp. 9-10, Besner 2007, p. 6).

3.1 Problems

Due to the strong reliance on private investment and (as a consequence) on economic conditions, the development of secondary pedestrian networks was at times stopped prematurely. This led to the problem of underused 'crippled' branches of what was supposed to be an integrated network. These fragments not only fail to provide the desired linkages on an urban scale. They also suffer from a lack of safety, social 'self-control' and vitality. This, in turn, impacts the frequency of pedestrian traffic and, thus, the economic basis of shops lining the bridges and tunnels. However, these shops are needed to finance the system in the first place (El-Geneidy et al 2011, pp. 39-43, City of Toronto CPD 2012, pp. 14).

A second problem faced by cities developing such secondary networks is the unbalanced use of (at least) two alternatives for pedestrians to reach a destination (usually either public sidewalks or private tunnels/bridges). By encouraging investors to contribute to the secondary network, the city runs the risk of supporting a system that does not complement the existing sidewalk-crosswalk-network, but rather renders it obsolete, unsafe, underused, and thus misses the city's original intention of increasing the sidewalk's capacity instead of replacing it (City of Toronto CPD 2012, appendix A-3).

As a way to overcome both potential dangers, we propose the restatement and refinement of the bonus systems: Instead of evenly rewarding every extension of the favored network customized bonuses could be offered for each individual block. This adaptation is facilitated by using Big Data and agent-based methods based on a regularly updated (quantitative and qualitative) survey and Big Data analysis of the pedestrian flows. As soon as new buildings get added, the networks' state would get updated, re-evaluated and the individual bonuses (if necessary) adapted. Instead of simply supporting a single network's growth, the refined bonus systems would thus encourage a balanced use of both the public and the private networks. So over a period of time, this now flexible bonus-system could even shift from rewarding tunnels or bridges to favoring the older ground-level system.

3.2 First part of the solution: Big Data

Any refinement of the bonus-system requires a broad political consensus. To be acceptable to both the public and the private sector it has to be based on detailed empirical data rather than on a bias towards the public (like the City of Montreal's current reading of the tunnels solely as an extension of the public transport systems, Besner 2007, pp. 3-4) or a dominance of the private business-community's interests (like in Toronto's Pedestrian Network's Master

Plan, City of Toronto CPD 2012, pp.7-11). By using already existing data-sets from sources as diverse as the real estate market, city polls and local police departments, underused and abandoned as well as overused and congested parts of each network can be localized and their degree of success or malfunction identified.

Applying Big Data methods, a meta study could be generated that by integrating various sources of the *status quo* could provide a thorough, coherent and unified mapping of both networks' pedestrian flows. Eventually, an 'actual ratio' of the existing flows at ground and at the secondary level could be provided for each individual building block.

Based on the survey's identification of 'hot' and 'blind' spots on both levels (i.e., points of successful and problematic development) these 'actual ratios' can be complemented by 'ideal ratios,' expressing the desired pedestrians' frequency for maintaining a balanced use and continuous growth of the networks. The resulting planning document could serve as a 'multi-level master plan' (i.e., a city map on two floors), anchoring each block within both networks, which only together provide the capacity for the necessary pedestrian movements.

3.3 Second part of the solution: Agent-based Modeling

Due to the optional, voluntary nature of bonus systems, the building's compliance to the 'ideal ratios' can only be encouraged, not enforced. As the development of a building project is influenced by a plethora of constraints (functional, typological, technical and economic) the outcome of the planning process is likely to differ from the prescribed ideal. Moreover, as even an approved project takes years to get realized, during construction its influence on the *status quo* and thus the 'ideal ratios' of adjacent building-blocks can only be guessed.

Building on research using agent-based modeling for analyzing and predicting pedestrian behavior (Batty et al 1998, pp. 32, 45, 56, Dijkstra et al 2012, p. 255, Zachariadis 2005, pp. 12-15), a simulation of future pedestrian flows could provide a means to compensate for this temporarily lack of empirical data (the time-span from the project's approval to its opening).

The predicted flows then provide a new basis for an updated master plan, for updated 'ideal ratios' and for the possible bonuses to be obtained by projects to come. Once available, empirical data of the finished project can be used to examine the agent-based model's various predictions and thus the reliability of its assumptions. Furthermore, the analysis can be continued through the life time of the project in order to make adjustments in places where predictions deviate from actual flows. While evaluating empirical data can help to render the networks' (agent-based) model more and more reliable, reality's constant feedback will additionally keep the master plan in touch with changes that lie outside its original parameters, but nonetheless may influence the total volume of pedestrian flows (e.g., economic booms, demographic shifts, climatic changes).

3.4 A Big Data challenge: Data availability

The above sketched sequence relies on the use of data that are either

1. publicly available (buildings and their links),
2. potentially available but not yet mapped (existing pedestrian flows), or
3. generated by models (the model's agents' pedestrian flows).

A quantitative survey of (publicly available) data deals with the networks' real-estate aspect (data required according to point 1 above). Within the area covered by the bonus system the built status of each block is rather transparent. i.e., it is known whether a block is linked to the secondary pedestrian network, to underground or other rapid transit stations, how much retail space is attached and whether the upper floors are used for office, housing, public programs or even more retail (e.g., Montreal, see El-Geneidy et al 2011, p. 43). A both quantitative and qualitative evaluation and modeling of the actual pedestrians' flow (in order to collect and generate data required according to points 2 and 3 above) represents more of a challenge. Various methods and models have been developed during the last 20 years to describe specific modes of city use by pedestrians including sociological, commercial, environmental and architectural aspects. The quantitative aspect—*pedestrian volume modeling*—is probably the best researched one, having been applied to a variety of buildings and urban areas (Raford & Ragland 2005, pp. 5-6, Urbitran Ass. 2004, p. 21, Timmerman 2009, p.xi, DCP NYC 2000, pp. 41-46, Boisvert 2005, pp. 10-15). (Surprisingly, this research has only focused on either public or private pedestrian networks; and has thus either addressed residential aspects or those of the business community, but not the *integration of both*.)

Encouraged by the current Big Data discussion, a next step could head for *ultimate complexity*, integrating as much information available as possible, in particular, combining the variety of well-established research threads with the opportunities current data mining can offer. This ambitious 'next step' would have to structure a broad data-basis, integrating the thoroughly studied influences of different (even marginal) conditions that guide the pedestrians' itineraries. By combining data provided by private and public sources using WiFi-Trace (Danalet et al. 2013, pp. 1-2), or publically installed 'sensor-nodes' (Mitchum 2014), not only the number of pedestrians on the various levels could be modeled, but also their itineraries, intentions and goals. E.g., in a survey of pedestrians conducted in Montréal in 1989 a variety of reasons for using the secondary network were revealed: work (51%), shopping (31%), entertainment and recreation (18%) (e.g., see Besner 1991, p. 13). Thus the desired balanced use of the pedestrian networks could also take into consideration qualitative aspects in addition to quantitative demands.

4. Conceptual case study 2: Optimizing current office designs

After the cellular and the open plan office and their integration in the so-called combi-office, office layouts have developed towards highly flexible workshop-like floor plans that comprise areas with call-center-like density of work spaces as well as lounge-like recreational areas. These work zones are typically populated by 'office nomads' who instead of occupying a dedicated work space now share desks or even work at home (cf. the Vitra furniture company's 'Citizen-Office-concept,' Vitra 2014). The most prominent examples of this development are to be found in the widely published office designs of Yahoo, Google, Microsoft, Sun, Cisco, Skype, Twitter, Facebook, and similar IT companies (Dainis 2014). Much more than simply saving floor area, the various open office concepts seem to be about adapting traditional office environments to the major changes in communication patterns brought about by the Internet, cell phones and various kinds of social media.

After a first phase of euphoria, spearheaded by companies from 'creative industries', even more traditional companies (including the US Department of Commerce, *ibid.* 2013, sect. 3) have realized that precious inner-company information is more and more traded in an informal way. Ad hoc meetings by the kitchen counter, in the temporary privacy of the elevator, or in a silent niche next to the copy machine were no longer considered a waste of time, but were effectively treated as efficient ways of supporting communication among co-workers.

4.1 Problems

As even the forerunners of the open office started re-evaluating their plans and rethinking the optimization of office spaces (cf., CEO Marissa Mayer's widely published end to Yahoo's work-at-home policy a year ago), the main challenge for more traditionally structured companies implementing these concepts was that realizing an open office concept could turn out to be a rather risky and expensive experiment. For one, transforming traditional office space into a highly flexible work-environment takes extra investments. To allow for more communicative, creative and thus productive performance in a space that can be divided up into densely packed and rather open areas and rearranged on a weekly basis, one has to provide movable furniture, flexible partition-walls, adjustable conference equipment, and easily adaptable air-conditioning; and it also requires extra manpower to manage the constant flow of people and project groups.

The worries are clear: what if all the remodeling does not pay off? What if only gossiping instead of communication is increased? What if the dynamic flow of people, work spaces and information renders ordinary work patterns impossible? (Cf., the rather

pessimistic findings of Waber et al. 2014, Smith 2013, p. 578, Pejtersen et al. 2011, p. 376, CABE 2005, pp. 43-54, Brennan et al. 2002, pp. 290 and Roelofson 2002, pp. 247).

These questions put a lot of responsibility both on the clients' and their architects' shoulders as they have to make sure that their design recipes will ultimately be successful from a business perspective. One way to put these assumptions on a more solid basis could be to employ Big Data analysis methods of the precedents combined with testing new designs using simulations of agent-based models.

4.2 First part of the solution: Big Data

A Big Data analysis that can provide a reliable basis for determining an office's size, shape and inner organization should cover as wide a range of information as possible. Hence, the precedents' evaluation should include the companies' national status and location, the extent to which the different office-concepts (i.e., open office, desk sharing, tele-working, etc.) are applied, and eventually make use of data depicting the employee's performance and job satisfaction (i.e., data usually collected by departments such as human-resources, facility-management, IT, R&D and the company's own security-systems).

An almost microscopic (although privacy-wise sensitive) perspective could be assumed by tracking the employees' movements throughout a working day: one could deduce the relative attraction of (data generating) spots like lobbies, meeting rooms, printers, the cafeteria's vending machines and, of course, the dedicated work-space. And by relating the employees' location to their use of various communication tools, one could additionally deduce each location's potential to trigger communication and trading of information (for the successful use of 'tracking badges', worn by employees' of the Norwegian 'Telenor' company on an opt-in basis see Waber 2014, p. 6).

As an architectural result of this (Big Data based) meta study the average amount of *floor area per workplace* can be approximated (a number that balances the optimization of space available with the employees' wellbeing) as well as the *floor area needed for service spaces* (whether and how these two numbers actually correlate would be an enlightening side-result of this study). These benchmarks, once established and accepted, could then play a decisive role in the process of finding an appropriate office long before a detailed planning of the office layout has even started.

Ideally, an analysis of open office parameters and company performance would establish the 'right' number of fixed and flexible work spaces, recreational space and amenities provided, and overall the percentage of space exclusively dedicated to circulation, storage and services. Such an analysis could complement current planning assumptions with a new corporate typology of contemporary generic office space and could help to replace 'office myths' by a corporate planning manual (for first efforts "to synthesize data collected in

workplace consultancy projects in order to create a decision-making tool for clients to better manage space usage in commercial buildings,” see Spacelab & SSL 2014, but also the internal research program of the planning and consulting company 'Gensler', Jerde 2013).

4.3 Second part of the solution: Agent-based Modeling

As a result, the client's assignment (usually a fixed number of departments and work spaces using the available floor area distributed throughout a given number of floors within a given building envelope) could be refined following the 'new typology': possible future employee behaviors can be evaluated in agent-based simulations of possible architectural setups (i.e., three-dimensional environments with specific layouts of circulation nodes, departments, and individual office spaces based on the dynamics of employee flows). These agent-based models could bring together an (economically computable, thus limited) amount of agents with an (also limited) variety of locations (differing in attraction) spread throughout a specific building. Given a number of 'encounters' the agents need to have in a given sequence, each agent randomly passes potential nodes of interaction, where the agent is rewarded with extra points, thus speeding up the entire office's reaching of the common goal. Hence, rather than only comparing the agent's movement patterns, the scoring function will bring to the fore potential flaws and strengths of different layouts for meeting and interaction among individuals (i.e., whether a certain arrangement leads to dead ends, detours or shortcuts). (For an outlook on the combinations of ABM and architectural planning see: Gao & Gu, 2009, pp. 170-171, and Kavulya, 2011, p. 694.)

4.4 A Big Data challenge: Data availability

We are quite aware that the data needed for the above sketched research methodology might not always be readily available, and in some cases it might be impossible to obtain. While data is indeed generated and collected both at the company's own discretion (as soon as their secret of success is subject of research) and within the constraints of the legal system, the use of employees' individual data poses major obstacles. Establishing a rigid and reliable 'firewall' of anonymity for both sources combined with the outlook of new (probably promising) insights in the company's inner workings will hopefully be an attractive enough incentive for a number of open offices to participate (cf. 'Remote Utilities' - blog's overview of the current legal situation considering employee monitoring and surveillance, RU, 2013). Moreover, as notions of privacy are shifting (e.g., with Millennials said to be much less concerned about revealing personal information publically), it might become easier for companies to obtain relevant data either by directly obtaining their employees' consent or by using information their employees already voluntarily provided on social media (how this kind

of data can then be integrated into ABMs or any kind of simulation models is a topic of current research in network science).

5. Discussion

The two conceptual case studies are examples of a variety of possible application scenarios for Big Data methods and agent-based models in urban design and architectural design. Specifically, the analyses point at the potential ways in which Big Data methods, either alone or in conjunction with agent-based models, can be of help to urban planners and architects. While each of the conceptual case studies was targeted at different problems at a different scale (urban neighborhood design vs. office design), there are important commonalities: both examples showed the potential of agent-based models to inform and optimize designs, and both examples also made the case for the application of Big Data methods, at the very least for mining the vast data that simulations of agent-based models can produce. In both case studies, agent-based models are able to reveal important temporal dynamics of human movements (in the first case study, the pedestrian flows and the way they are impacted by changes in the layouts of the pedestrian networks; in the second case, the trajectories of employees through the office spaces and how they are impacted by changes in the functional roles of spaces).

Depending on the complexity of the set of agent rules that govern the behavior of individual agents (e.g., the extent and level to which preferences, goals, interaction dynamics, etc. are explicitly modeled), the results might require significant data mining efforts and statistical analysis to expose relationships among the independent and dependent variables. Some of these might be rather obvious and expected (e.g., the extent to which street-level pedestrian crossing is preferred to a pedestrian bridge or that using a tunnel is likely influenced by an individual's goal of avoiding steps, or how the frequent need for coffee could result in increased communications if the spatial layout of the cafe or coffee-making area in an office environment supports lingering). Others will be much less obvious (e.g., the extent to which cognitive factors influence people's perceptions of spaces and designs and thus lead to alterations in their behavior, see also recent attempts at isolating such factors in Sussman & Hollander 2015). And yet others will be much more complex to obtain, which is why applying Big Data methods to the simulation results will likely lead to surprising, yet potentially very useful findings that designers did not consider in the past, simply because the relationship is only revealed through analyses of large enough data sets.

6. Conclusion

In this chapter we argued that agent-based modeling in conjunction with Big Data analysis methods could prove invaluable to urban designers and architects and we used two

conceptual case studies to analyze where agent-based models in conjunction with Big Data analysis methods could specifically improve urban and architectural designs. By using not just historical data (as is typically the case in urban computing), but exploring patterns in data generated by agent-based simulation models, new designs can be explored in unprecedented ways, taking larger numbers of factors into account than ever before. The difference to current orthodoxy in design is clear: rather than relying on what we know from past experience, we can now generate new experience in hypothetical situations that will allow us to avoid pitfalls (designs that do not work) and to discover solutions (designs that do work) without having to wait for a design to be evaluated in reality. The next step now is to put the analyses to work in concrete design settings to verify that our conceptual analyses hold up to empirical litmus test.

References

- Saad Ali, Ko Nishino, Dinesh Manocha, Mubarak Shah (2013). *Modeling, Simulation and Visual Analysis of Crowds: A Multidisciplinary Perspective*, Springer Science & Business Media, 422 pages.
- M. Batty, B. Jiang, and M. Thurstain-Goodwin (1998). "Local movement: agent-based models of pedestrian flows". *CASA Working Papers. Centre for Advanced Spatial Analysis* (UCL) London, UK
<http://eprints.ucl.ac.uk/225/>
- Jacques Besner (1991). "Historical Perspective" and "A Model of Partnership?", *Actualité immobilière, Special issue: Montréal Underground*, UQAM, Montréal, 1991, 3-11, 12-24
- Jacques Besner (2007). "Develop the Underground Space with a Master Plan or Incentives. Associated research Centers for the Urban Underground Space", *Proceedings fo 11th ACUUS Conference*, Athens, Greece 2007, 1-7
http://observatoiredelavilleinterieure.ca/documents/ACUUS_XI-Besner.pdf
- Trevor Boddy (1992), "Underground and Overhead: Building the Analogous City," *Variations on a Theme Park: The New American City and the End of Public Space*. ed. Michael Sorkin, New York, 1992, 123-153
- Michel Boisvert (2005). "Modeling pedestrian flows in Montreal's indoor city". Associated research Centers for the Urban Underground Space, *Proceedings of 10th ACUUS Conference*, Moscow, Russia, 1-24
http://www.observatoiredelavilleinterieure.ca/documents/ovi_moscow_2005.pdf
- Aoife Brennan, Jasdeep S. Chugh, and Theresa Kline (2002). "Traditional versus Open Office Design: A Longitudinal Field Study". *Environment and Behavior*, 2002, 34, 279-299
<http://senate.ucsf.edu/2013-2014/mb4-brennan%20et%20al%20article%20on%20moving%20into%20open%20space%20offices.pdf>
- CABE-Commission for Architecture and the Built Environment and British Council for Offices (2005). "The impact of office design on business performance". *CABE-Homepage*, 1-79, London
<http://webarchive.nationalarchives.gov.uk/20110118095356/http://www.cabe.org.uk/publications/the-impact-of-office-design-on-business-performance>
- CCD, Urban Center for Computation and Data (2015)
www.urbanCCD.org
- City of Toronto, City Planning Division (2012). "PATH Pedestrian Network, Master Plan". *City of Toronto, Ontario*
http://www1.toronto.ca/city_of_toronto/city_planning/transportation_planning/files/pdf/path_masterpla_n27jan12.pdf
- Margaret Crawford (1992), "The World in a Shopping Mall," *Variations on a Theme Park: The New American City and the End of Public Space*. ed. Michael Sorkin, New York, 1992) 3-30.

Jianqiang Cui, Andrew Allan, and Dong Lin (2011). "Influencing Factors for Developing Underground Pedestrian Systems in Cities". *Proceedings of Australasian Transport Research Forum 2011*, 28 - 30 September 2011, Adelaide, Australia

http://www.atrf.info/papers/2011/2011_Cui_Allan_Lin.pdf

Jianqiang Cui, Andrew Allan, and Dong Lin (2013). "The development of grade separation pedestrian system: A review". *Tunnelling and Underground Space Technology*, 38, September 2013, 151–160

<http://research-hub.griffith.edu.au/display/nf8e897b0880aa4ade00e0d3e9a69251f>

The Chicago Architecture Foundation (2013). "Chicago: City of Big data Exhibition". *Official Exhibition-Homepage*, Chicago

<http://bigdata.architecture.org/data-and-design/>

Fabrício F. Costa (2012). "Big Data in Genomics: Challenges and Solutions(Is Life Sciences Prepared for a Big Data Revolution)?" *G.I.T. Laboratory Journal* 11-12

Bill Cull (2013). "3 Ways Big Data is Transforming Government". FCW

<http://fcw.com/articles/2013/09/25/big-data-transform-government.aspx>

CUSP, Center for Urban Science and Progress, NYU (2015)

<http://cusp.nyu.edu/>

Dainis c/o hongkiat.com (2014). "Creative & Modern Office Designs Around the World".

Hongkiat design and technology weblog (hongkiat.com)

<http://www.hongkiat.com/blog/creative-modern-office-designs/>

Antonin Danalet, Bilal Farooq, Michel Bierlaire (2013). "Towards an activity-based model for pedestrian facilities Swiss Transport Research Conference", Proceedings of 13th STRC Conference, Ascona, 1-31

http://www.strc.ch/conferences/2013/Danalet_Farooq_STRC13.pdf

DCP NYC; Department of City Planning City of New York (2000). "Midtown Manhattan Pedestrian Network Development Project." *Official Website of the City of New York*

http://www.nyc.gov/html/dcp/pdf/transportation/mmp1_full.pdf

DCP NYC; Department of City Planning City of New York (2014). "Privately Owned Public Spaces / history". Official Website of the City of New York

http://www.nyc.gov/html/dcp/html/pops/pops_history.shtml

Jan Dijkstra, Harry Timmermans, and Bauke de Vries (2012). "Activation of Shopping Pedestrian Agents— Empirical Estimation Results". *Applied Spatial Analysis and Policy*, 6, 4, 255-266, Springer Netherlands

<http://link.springer.com/article/10.1007%2Fs12061-012-9082-3>.

[Charles Duhigg](#) (2012). "How Companies Learn Your Secrets", in New York Times Sunday Magazine, February 19, 2012, p. MM30.

<http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html>

Ahmed El-Geneidy, Lisa Kastelberger, and Hatem T. Abdelhamid (2011). "Montreal's roots, Exploring the growth of Montreal's Indoor City". *The Journal of Transport and Land Use*, University of Minnesota, 4, 2, Summer 2011, 33-46

http://tram.mcgill.ca/Research/Publications/Montreal_indoor_city_final_in_JTLU.pdf

Amr Gandomi and Murtaza Haider (2015). "Beyond the hype: Big data concepts, methods, and analytics." *International Journal of Information Management*, 35, 2, April 2015, 137-144.

Yin Gao and Ning Gu (2009). "Complexity, Human Agents, and Architectural Design: A Computational Framework". *Design Principles and Practices*, 3, 6, 115-126.

<http://www.sciencedirect.com/science/article/pii/S2095263512000167>

Patrick Grim, Gary R. Mar, Paul St. Denis (1998). "The Philosophical Computer: Exploratory Essays in Philosophical Computer Modeling". Bradford, 333 pages.

Volker Grimm, Uta Berger, Finn Bastiansen, Sigrunn Eliassen, Vincent Ginot, Jarl Giske, John Goss- Custard, Tamara Grand, Simone K. Heinz, Geir Huse, Andreas Huth, Jane U. Jepsen, Christian Joergensen, Wolf M. Mooij, Birgit Mueller, Guy Pe'er, Cyril Piou, Steven F. Railsback, Andrew M. Robbins, Martha M. Robbins, Eva Rossmannith, Nadja Rueger, Espen Strand, Sami Souissi, Richard A. Stillman, Rune Vabo3 Ute Visser, Donald L. DeAngelis (2006). "A standard protocol for describing individual-based and agent-based models". *Ecological modelling* 198, 115-126.

Chris Jerde (2013). "Design Value Through the Lens of Big Data". Homepage Gensler(gensler.com): Gensler on Work

<http://www.gensler.com/work/2013M/26/design-value-through-the-lens-of-big-data.html>

Geoffrey Kavulya, David, Burcin, and Becerik-Gerber (2011). "Designing in complex system interaction: Multi-agent based systems for early design decision making". *Proceedings of the 28th ISARC International Association for Automation and Robotics in Construction*, Seoul, Korea, 694-698.

<http://www.iaarc.org/publications/fulltext/S21-2.pdf>

Fosco Lucarelli (2014). "The Endless Interior: Calgary's Plus 15 Skywalk System". *Socks-studio.com online-magazine*, Paris

<http://socks-studio.com/2014/04/13/the-endless-interior-calgarys-plus-15-skywalk-system/>

Viktor Mayer-Schönberger and Kenneth Cukier (2013). *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Eamon Dolan/Houghton Mifflin Harcourt

Rob Mitchum (2014). "From Spreadsheets to Solutions: New Platform Enables Next Generation of Open City Data". *University of Chicago-Homepage*, Chicago (www.uchicago.edu); Articles 2014

<https://www.ci.uchicago.edu/press-releases/spreadsheets-solutions-new-platform-enables-next-generation-open-city-data>

Aaron Moore (2013). "Trading Density for Benefits: Section 37 Agreements in Toronto"

IMFG Perspectives", *Toronto*, 2, 2013, 2-7

http://munkschool.utoronto.ca/imfg/uploads/221/imfg_perspectives___moore_%28feb_2013%29.pdf

MicroSoft Urban Computing (2015)

<http://research.microsoft.com/en-us/projects/urbancomputing/>

Noah Raford and David Ragland (2005). "Pedestrian Volume Modeling for Traffic Safety and Exposure Analysis". *Transportation Research Board, Washington*, TRB 85th Annual Meeting Compendium of Papers CD-ROM

<https://escholarship.org/uc/item/9cn8d3nq>

one simulation.com (2014). "Evacuating of people". *company-homepage*, Leiden, NL

http://www.onesimulations.com/index.php?p=fire_safety

one simulation.com (2014). "Building exodus". *Elements-Exhibition*, Biennale, Venice

<http://www.worldarchitects.com/images/CmsPageElementImage/40/31/87/5395834fdf70407a934b4cd30ab56af9/5395834fdf70407a934b4cd30ab56af9.jpg>

Jan H Pejtersen, Helene Feveile, Karl B Christensen, and Hermann Burr (2011). "Sickness absence associated with shared and open-plan offices". *Scandinavian Journal of Work, Environment & Health*, 37, 5 (September 2011), 376-382

<http://www.istor.org/discover/10.2307/23064898?uid=3737528&uid=2&uid=4&sid=21104696972827>

Remote Utilities (2013). "Employee Monitoring and Surveillance: Important Laws You Should Know Remote Utilities Blog", *Usoris Systems LLC*, Seychelles

<http://www.remoteutilities.com/about/blog/Remote.Utilities/employee-monitoring-and-surveillance-important-laws-you-should-know/>

Paul Roelofsen (2002). "The impact of office environments on employee performance".

Journal of Facilities Management, 1, 3, 247-264

<http://www.emeraldinsight.com/doi/abs/10.1108/14725960310807944>

Scheutz, M. and Harris, J. (2011). "An Overview of the SimWorld Agent-based Grid Experimentation System". In Dubitzky Werner, Krzysztof Kurowski, and Bernard Schott (eds.) *Large- Scale Computing Techniques for Complex System Simulations*, Wiley

Alan D. Smith (2013). "Online social networking and office environmental factors that affect worker productivity". *International Journal of Procurement Management*, 6, 5, 578-608

<http://www.inderscience.com/info/inarticle.php?artid=56173>

William Solecki, Karen C. Seto and Peter J. Marcotullio (2013). "It's Time for an Urbanization Science", *Environment Magazine*, January-February

<http://www.environmentmagazine.org/Archives/Back%20Issues/2013/January-February%202013/urbanization-full.html>

Michael Sorkin (Ed.) (1992). "Variations on a Theme Park". *Farrar, Straus and Giroux*, New York

http://books.google.at/books/about/Variations_on_a_Theme_Park.html?id=QMhohDJgHIYC&redir_esc=y

Spacelab, Space Syntax Laboratory at The Bartlett, Centre for Advanced Spatial Analysis (2014).

"Knowledge Transfer Partnership project 'Big Data in the Office'". *Space-Syntax-Homepage*

https://www.bartlett.ucl.ac.uk/space-syntax/research/projects/ktp_big_data_in_the_office

Ann Sussman and Justin Hollander (2015). *Cognitive Architecture: Designing for How We Respond to the Built Environment*, Routledge.

Harry Timmermans (Ed.) (2009). "Pedestrian Behavior: Models, Data Collection and Applications." *Emerald Group Publishing*, Bingley UK

Torrens, P.M. & McDaniel, A. (2013) "Modeling geographic behavior in riotous crowds". *Annals of the Association of American Geographers*, 103(1): 20-46

Urban Redevelopment Authority (2014). "Central Area Underground Master Plan Revisions to the Cash Grant Incentive Scheme for Underground Pedestrian Links". *Official Website of the URA Singapore*

<http://www.ura.gov.sg/uol/circulars/2012/aug/dc12-12.aspx>

Urban Systems Collaborative (2015)

<http://urbansystemscollaborative.org/>

Urbitran Associates (2004). "Pedestrian Flow Modeling for Prototypical Maryland Cities". *Maryland Department of Transportation*, Hanover, MD

http://smartgrowth.umd.edu/assets/cliftondaviesallenraford_2004.pdf

U.S. Department of Commerce (2013). "Space Allowance and Management Program". *Official Website of United Department of Commerce*

http://www.osec.doc.gov/opog/dmp/daos/dao217_21.html

Vitra AG (2014). "The Citizen-Office-Concept", *Homepage Vitra-AG (vitra.com)*, Birsfelden, Switzerland

<http://www.vitra.com/en-br/office/index-concepts/citizenoffice>

Ben Waber, Jennifer Magnolfi, and Greg Lindsay (2014). "Workspaces That Move People". *Harvard Business Review*, October 2014, online-version

<http://hbr.org/2014/10/workspaces-that-move-people/ar/2>

Yanagisawa, D., Kimura, A., Tomoeda, A., Ryosuke, N., Suma, Y., Ohtsuka, K., Nishinari, K. (2009). "Introduction of frictional and turning function for pedestrian outflow with an obstacle". *Physical Review E* 80(3), 036110.

Vassileios Zachariadis (2005). "An Agent-Based Approach to the Simulation of Pedestrian Movement and Factors that Control it". *Proceedings of Computers in Urban Planning and Urban Management CUPUM 2005*, 1-16, London, UK

<http://128.40.111.250/cupum/searchpapers/papers/paper372.pdf>