Journal of Comparative Urban Law and Policy

Volume 5
Issue 1 *A Festschrift in Honor of Arthur C. Nelson on the Occasion of his Retirement - Agenda for Building a Changing World Responsibly: Commentaries and Reflections by Leaders in Urban Planning, Policy, and Design*

Article 32

From the Abacus to Big Data: The Evolution of Data-Driven Planning in the U.S. and Where the Field will be Headed

Keuntae Kim
University of Utah, kimprins@hotmail.com

Recommended Citation

Kim, Keuntae () "From the Abacus to Big Data: The Evolution of Data-Driven Planning in the U.S. and Where the Field will be Headed," *Journal of Comparative Urban Law and Policy*: Vol. 5 : Iss. 1 , Article 32, 401-424.
Available at: https://readingroom.law.gsu.edu/jculp/vol5/iss1/32

FROM THE ABACUS TO BIG DATA

THE EVOLUTION OF DATA-DRIVEN PLANNING IN THE U.S. AND WHERE THE FIELD WILL BE HEADED

Keuntae Kim[*]

ABSTRACT

The nature of planning involves a set of decision-making processes to fulfill people's needs and expectations of where they live, work, and play. Dealing with the nature of planning—complexity, uncertainty, and disagreement—requires specific tools to explore various aspects of the built environment as a whole. Various types of data have been extracted, transformed, and loaded to describe the past and current conditions of the built environment, and planners have developed and applied data-driven planning tools to explore the knowns and unknowns of the urban futures and transform them into a set of actions based on the goals with consensus. This article identifies the evolution of systematic ways of collecting data, setting up criteria, and analyzing them according to the contextual features of planning tools, focusing on where the field is headed for planning.

INTRODUCTION

Since the beginning of the twentieth century, urban planning professionals have developed many planning tools and systems to envision urban futures in a comprehensive, transparent, and effective way. Throughout planning history, the evolution of planning tools depends on data availability and technologies that help planners build models for planning issues and types of urban problems and solutions that planners like to seek (Silva and Wu, 2012; Klosterman 1997). Outcomes estimated from various analysis models have been frequently used as additional data for further modeling or analyses on other complex urban planning problems. Big data analytics, machine learning, and deep learning algorithms are adopted from computer science to urban planning, which now enables planners to apply these techniques to devise more comprehensive and accurate planning tools for predicting a range of possible urban futures, testing them, and selecting the most desirable planning alternatives as a comprehensive plan—that is, data-driven scenario planning. Ultimately, scenario planning outcomes drawn from scenario planning tools have played a major role in generating additional data for further analysis to reduce the complexity and uncertainty of urban futures.

---

[*] Research assistant, Metropolitan Research Center, University of Utah. Doctoral degree awarded in 2021.

This article aims to provide planners with some insights into how reasoning and methods of tools in planning have been evolved in terms of the development of data processing and information technology. Considering existing literature about the development of tools in planning, this paper will attempt to expand the historical overview of data-driven tools into the era of the emergence of modern planning at the beginning of the twentieth century. Also, this paper will try to integrate existing literature about tools in planning into one piece of the overarching overview of data-driven tools in planning. In planning literature about data-driven tools, studies suggest an overview of computer-aided planning tools based on the development of computer technology and geo-information systems, whereas some literature review studies explain the development of planning tools or models separate from advances in data processing (Klosterman, 1997; Harris, 1999; Harris and Batty, 1993; Klosterman, 1999b; Guhathakurta, 1999). There is still little research that attempts to explain an overarching review of the history of data-driven planning tool development (LeGates, Tate, and Kingston, 2009), but it does not mention if there were any development of data-driven tools or models when they were seriously attacked by planners or shift of paradigm of planning occurred against the evolution of these tools.

From these continuum perspectives, this paper overviews the evolution of data-driven planning tools according to the similar terms used by urban professionals over different periods to represent the concept and framework of data-driven tools in planning. This paper is structured as follows: after the introduction section, five distinct development phases and one shift identified in this study are explained in chronological order, and key terms that stand for data-driven planning processes are also mentioned in each section. Then, this paper explains gaps that occurred between development phases and summarizes how data-driven planning tools were evolved over these periods. Recent challenges of data-driven tools in planning—smart grid, smart city digital twin, digital divide, autonomous vehicle, the Internet of things (IoT), and contact tracing in the COVID-19 pandemic—are briefly covered by linking them with the current data-driven planning processes. The conclusion section summarizes all findings of this study.

PHASE ONE: EMERGENCE OF SCIENTIFIC PLANNING (1909 – 1950)

It is still controversial to define when modern data analysis tools or models in urban planning emerged, but the U.S. National Conference on City Planning held in 1909 was considered as the earliest time when modern data-driven planning tools or methods were first identified in the urban planning practices (LeGates et al. 2009). As a primitive form of data-driven planning tools, scientific planning was considered as a "formal" method, compared to the "picturesque" methodological approach found in the City Beautiful Movement (Nolen, 1909). In particular, Ford's assertion of scientific planning at the fifth U.S. National Conference on City

Planning held in Boston mentions that, like other industries and engineering fields, urban problems such as convenience, health, and efficacy can be addressed as scientific subjects (Ford, 1913). Urban planners in this period considered a city as a big laboratory full of complexity. Scientific planning during this phase aimed at remodeling cities to reduce currently bad conditions of cities and transform them into the ideal condition through “experimental design, hypothesis testing, and the interpretation of data for exploring, predicting, and controlling empirical phenomena in a rational manner (Nolen, 1909; Kim, 2013).”

Dominated as a primary urban planning approach between the late 1910s and the 1950s, scientific planning regarded city planning as what Ford (1913) called “as definite a science as pure engineering.” Most efforts of scientific planners were oriented toward setting up an experiment and collecting data suitable for it. Simple observation and surveys with a quantifiable questionnaire were mostly used to build up a dataset suitable for quantitative analysis. Data collected through these methodologies had descriptive features that represented simple and structured information within a larger picture of the urban conditions—for example, population, street, buildings, water supply, housing, etc. To avoid data collection biases, data built during this period were concentrated on gathering physical dimensions of the built environment. During this period, a key issue of data management was to accumulate data in the urban planning system and think about how to turn them into knowledge useful for analysis of the existing urban conditions (Forsyth, 2012).

Nevertheless, structured racism during this period affected data collection and analysis for planning experiments, and analysis results were used to strengthen inequality in urban development patterns. For instance, the 1930 and 1940 decennial Census data of population and housing were collected through “mailed-based” survey, and lack of survey literacy among Black and other immigrant populations led to biased data collection even if the overall socioeconomic inequality and population turnovers could be identified from the Census data, which were used to differentiate neighborhoods by race, ethnic, and class (Turner 1954; Green 2015; Margo 1986; Connor & Storper 2020). The federal government’s Home Owners’ Loan Corporation (HOLC) maps of redlined neighborhoods made data quality biased and unreliable for planning. Lack of data management and analysis techniques and exploratory data analysis using simple aggregation of data at the larger spatial scale did not work very well in estimating the future growth or development during this period.

Although scientific planning was the first structured attempt to accumulate data as a stock within an urban system, it was not effective in establishing a set of processes or models that can explain the urban conditions or phenomena as a whole. In the scientific planning approach, planners set up an experiment with many other influential factors controlled through assumptions or rules based on their intuition or some similar planning cases they are familiar with. Scientific planning uses univariate or bivariate descriptive analyses to produce outcomes through one or a few mathematical equations. These outputs can be used only to find one answer to one simple and structured urban planning problem. Data in scientific planning were not accessible and shared only by planning elites (Nolen, 1909; Ford, 1913; Forsyth, 2012; LeGates, Tate, and Kingston, 2009).

In this phase, federal-level workforce housing and community reform projects were typical examples of applying scientific planning approaches to planning (Fairfield, 1994; Sagasti, 1973; Topalov, 1990). To provide housing units for veterans returning to the U.S. after World War I, the U.S. federal government applied the scientific planning process to building suburban residential development projects. These suburban residential projects were considered a set of "model villages or communities based on working-class status." Different types of housing units were provided to workers based on their skills and position in the industrial mass production process (Topalov, 1990). Data used in the projects were physical dimensions of different types of housing units and layouts of communities, and planners believed that model villages or communities could achieve both economic efficacy and social efficacy through the aggregation of these physical dimensions and universal standards of housing units. In scientific planning, layouts of communities and villages were simply an aggregation of different types of housing units. A key mathematical equation for analysis was based on the 'supply and demand' or 'capacity models' from economics and pure engineering fields (Topalov, 1990; Fairfield, 1994; Sagasti, 1973).

PHASE TWO: URBAN MODELS AS DATA-DRIVEN AND SYSTEMATIC TOOLS (1950 – 1973)

As additional data were available, it became evident that simple univariate and bivariate data analysis could not address complex urban problems and explain interactions among various aspects of the built environment as a whole. Particularly, as planners began to recognize "planning as applied science" that put more emphasis on consultant approaches to considering more commercial and readymade solutions, more systematic approaches to analyzing what urban problems were needed to address for more efficient land use development in the future (Madanipour, 2010; Klosterman, 1997). Also, outputs of scientific planning analyses were drawn based on the ideal, utopian basis and far from the reality and planning practice.

To fill this gap between ideal planning experiments and planning practices, planners began to think about developing "systematic and interactive approaches" to finding more comprehensive solutions to urban problems—that is, urban models. Urban models are defined as tools that allow both planners to understand "the basic activities of urban areas as a major part of the scenario (Pack, 1975).” In this phase, more datasets and knowledge of data processing and management contributed to analyzing one planning phenomenon in urban areas that is more hierarchical than one fragmented and specific problem-solving process in scientific planning. Planners were able to examine possible impacts of one planning phenomenon on the future urban growth and development, allowing planners to conduct the comprehensive land-use analysis through large-scale urban models (Lee, 1994). Also, in academia, planning professors and students studied the potentials of urban models as tools to comprehend relationships among existing aggregated datasets and analyze more hierarchical and complicated urban problems—especially land-use planning from various perspectives—through a set of mathematical forms used for urban models (Lee, 1994; Klosterman, 1994).

While having faith in quantitative analysis and experiments in the former period, planners used a computer or a handheld scientific calculator to model a wide range of the spatial scopes in the built environment—from a Census block to metropolitan areas or regions. Also, more data can be recorded and accumulated in a digitalized way, and the functionality of urban models was greatly improved through the embodiment of sets of mathematical equations that represented several physical and economic aspects of an urban system (Batty, 1994; Guhathakurta, 1999). Therefore, in this period, urban models mean a large-scale model aided by personalized digital instruments that enabled planners to shift their paradigm from “planning as the art” in the late nineteenth century to “planning as pure and applied science” with strong confidence (LeGates, Tate, and Kingston, 2009; Klosterman, 1997; Forsyth, 2012; Batty, 1994; Harris and Batty, 1993).

Data collection and management in urban models over this period were generally focused on analyzing existing land use patterns and transportation volume in metropolitan areas to predict the future land use and transportation impact through urban models (Harris, 1965; Webber, 1959; Vorhees, 1959; Blumfeld, 1959; Hansen, 1959). For example, through analysis of the future land use patterns and transportation in the Chicago metropolitan area, Hamburg and Creighton (1959) developed an urban model of the future land use pattern through analyses of population and worker changes by distance from the central business district (CBD) area.

As with scientific planning in the former period, urban models have similar limitations. First, data collected and used to analyze land use and transportation were still concentrated on descriptive and scientific dimensions of the physical environment. During this period, advances in computer technology contributed to accumulating a large number of datasets for land use, transportation, housing, streets, water, and so on. Nevertheless, the lack of information and knowledge about linking unscientific values with scientific ones made the functionality of urban models incomplete. Second, the basic hypotheses of urban models still stick to the ideal and utopian growth of future urban development. Finally, despite computer technology, integrating many different variables into one single model to produce more comprehensive outcomes remains a key issue. These limitations led to critiques of urban models done by Douglas B. Lee in 1973, which will be covered in the next section.

THE FIRST SHIFT: CRITIQUES OF LARGE-SCALE URBAN MODELS AND MATURING DATA-DRIVEN TOOLS AND PROCESSES (1973 – 1980S)

Existing literature on critiques of large-scale urban models in the 1950s and the 1960s suggest that Lee's article "Requiem for Large-scale Models" in 1973 contributed to reconsidering performance and even usefulness of data-driven tools in planning by summarizing seven "sins" of urban models in that article (Klosterman, 1994; Batty, 1994; Guhathakurta, 1999; Forsyth, 2012). Also, the "four rough guidelines for the modeling movement" suggested by Lee contributed to proving former large-scale urban models and their outputs invalid in actual planning practices and theories. Following this harsh attack, data management and processing for developing urban models and analyzing urban problems in a scientific way seemed to disappear in urban planning. In fact, during the 1970s, there were no ways to prove his assertions of urban models incorrect by developing more innovative models due to the limited capacity of data processing and management and computer technology.

On the other hand, his arguments of urban models also played a role in re-thinking about maturing data-driven planning tools toward a constructive way. While criticizing Lee's assertions, Harris (1994) argues that advances in computer and data management technologies and statistical analytic methods such as non-linear equations, regression analysis, and equilibrium formula can make data-driven planning tools reflect the real-world urban conditions and phenomena within models. Also, two years after Lee's article, Pack (1975) conducted a wide range of a mail survey to about 1,500 planning agencies and found that one-third of them were still using urban models to make decisions about specific planning issues such as future land use and transportation pattern changes. Also, in planning practice, the extent of using urban models was expanding across various spatial levels. Planning agencies thought that through consultation with private companies, urban models had the potential to produce better outcomes and modeling processes in the future (Pack, 1975).

PHASE THREE: REVISITING DATA-DRIVEN TOOLS—DECISION SUPPORT SYSTEMS (LATE 1970S – EARLY 1990S)

After recognizing Lee's harsh realities about data-driven planning tools, advances in microcomputer technology in the late 1970s and the early 1980s greatly affected the development of data-driven tools in urban planning and provided planners with another opportunity to reconsider the development of data-driven tools in planning. Using microcomputers (i.e., Atari, Apple, IBM, HP, etc.) greatly increased individual ability to process a large amount of data for modeling the physical environment of a city and helped planners organize them at various spatial levels and planning conditions. Also, the spread of microcomputers enabled planners to share data. Finally, knowledge and skills for data building and management became one of the academic fields called informatics.

Along with the emergence of informatics as a new academic research field, the concept and term decision support system (DSS) emerged in the late 1970s and the early 1980s (Sprague, 1980; Bonczek, Holsapple, and Whinston, 1979; Klosterman and Landis, 1988). Considering a broad and technical definition, a decision support system means computerized and non-computerized tools that help system users make decisions for various problems. In urban planning, the concept and framework of decision support systems were first shown in planning journals and relevant academic field journals in the late 1970s and the early 1980s (Bonczek, Holsapple, and Whinston, 1979; Han and Kim, 1989; Klosterman, 1992; Harris, 1989; Klosterman and Landis, 1988). In a narrow definition, decision support systems can be defined as computer software with a set of analytic models inside the systems that allow planners to make decisions for various planning issues such as environmental analysis, land use planning decisions, location models for buildings and activities, development control, and so on (Klosterman, 1992). Also, some literature mentions that DSS is "a distinctive type of urban information system. (Han and Kim, 1989)."

To overcome the limits of former urban models raised by Lee (1973), the framework and principles of decision support systems are designed to find out solutions to poorly structured urban problems, such as estimating impacts of land-use changes on future urban growth (Han and Kim, 1989; Sprague, 1980). Also, decision support systems are designed and operated based on a set of "decision models," and these models produce outputs using relevant datasets to support decision-making. The user interface within the DSS system helps planners transfer datasets from various sources to decision models as raw data. While maintaining the qualitative analysis process of interpreting the results and developing implementation plans, the interaction between the DSS system and data played a major role in thinking of the data processing process as a more rational, scientific, and iterative process than the processes of urban models and scientific planning (Han and Kim, 1989).

Most of the planning literature during this period mention that the development of Geographic Information Systems (GIS) and geo-information databases contributed to creating more integrated modeling of the built environment (Rubenstein-Montano, 2000; Klosterman, 1992 and 1997; Han and Kim, 1989; Guhathakruta, 1999; Adler, 1987; Batty, 1988; Harris, 1989). From the data management perspective, DSS allowed planners to create thematic maps representing outcomes of DSS by combining databases with geo-information, which includes geometry information such as streets dimensions, lot areas, slope, and so on. In terms of data processing, the diversity of datasets contributed to making outputs of DSS close to reality and reducing the uncertainty of the plan proposals developed by DSS. Some literature defines DSS systems integrated with spatial information as Spatial Decision Support System (Geertman and Stillwell, 2004; Han and Kim, 1989; Klosterman 1994 and 1997).

The spread and use of DSS tools over this period contributed to making data-driven planning tools remain feasible. Nevertheless, in terms of data management and processing, some planning literature mentions that several limits exist in the DSS. First, most DSS software and tools are generally designed to find short-term decision-making or solutions to complex spatial problems (Geertman and Stillwell, 2004; Clarke, 1990). When DSS tools integrated GIS, they were able to analyze impacts of future urban changes in a more comprehensive and accurate way, but lack of "sufficient support to the development of long-term plans and objectives" made DSS tools less plausible for creating long-term alternatives (Harris, 1989). Second, even if geo-information datasets allowed planners to use them as additional datasets for accurate analysis and future estimation, most data used in DSS represent the physical properties of the built environment. This means that DSS did not include any models for unscientific values. Finally, concerning the first limitation, DSS software and tools were not effective in producing long-term alternatives as a whole. In particular, the last limitation of DSS software led to the development of the planning support system (PSS), which is covered in the next section.

PHASE FOUR: COMBINING GEOINFORMATION DATA WITH PLANNING TOOLS – PLANNING SUPPORT SYSTEM (1990S)

The concept of Planning Support System (PSS) first emerged in the late 1980s (Harris, 1989). As for the emergence of PSS, two contextual and one external reason can be identified: first, PSS was developed "in response to planners' fascination of GIS (Klosterman, 1999)." Second, as a new type of data-driven planning tools deviated from DSS, there were strong demand and movement among planners that more visualized and integrated planning support tools would be required to meet diverse needs and evaluate alternatives based on multiple criteria. Third, due to the prevalence of personal computers among planners and citizens, new roles of planners such as setting up rules, explaining the urban conditions to people, analyzing the impacts of the future urban changes, and making plans were needed (Harris, 1989).

In terms of the system structure, PSS has a lot in common with DSS or SDSS. It contains a set of models for estimating future impacts of alternatives. It creates and uses databases from various sources and spatial information from GIS, such as a map shapefile with physical properties of polygons. However, the main difference between DSS (or spatial decision support systems, SDSS) and PSS is that the latter puts more emphasis on producing long-term outputs in a highly visualized way and integrating outputs of planning elements such as land use and transportation together to get a more comprehensive and accurate result than the former (Harris, 1989; Geertman and Stillwell, 2004). Particularly, unlike DSS or SDSS, PSS is more focused on providing functions that can automatically estimate optimized planning alternatives through planning support models inside the system (Harris, 1989). These characteristics of PSS show that its technical approach to data management and processing is more intuitive than DSS and based on what planners prefer to evaluate alternatives.

In terms of the structural framework, PSS takes a more systematic approach, configuring models based on several general components. Hopkins (1999) explains that PSS is operated through a series of abstract elements rather than actual system elements—such as actors, activities, flows, facilities, regulations, etc. These elements can be grouped into three main components: the “objects” that determine urban development patterns and operate the tool for analysis; the “interface views” that allow planners to observe alternatives they create at various levels of “abstraction and completeness;” and “task and tools” that includes producing options, estimating the performance of each alternative, evaluating performance and validity of PSS tools, and communicative and collaborative planning process while operating a PSS tool.

Although planners devised and operated many PSS tools (such as TRANUS, SPARTACUS, METROPILUS, and so on) during this period, what is significant in PSS at that time was that it provided both theoretical and practical frameworks of how to make planning options generated from the tools more comprehensive and accurate according to goals and assumptions. As Harris (1994) pointed out, the emergence of PSS returned to large-scale urban models that Lee (1973) criticized.

PHASE FIVE: ENVISIONING THE FUTURE - SCENARIO PLANNING AND GEODESIGN (2000S – PRESENT)

This paper attempts to explain the terms scenario planning and GeoDesign in the same section because they both are emerging concepts and because tools are not yet widely shared among planners or do not exist. As a planning support system, the concept of scenario planning was first used by urban planners in the late 1990s and the early 2000s. Adopted from other academic fields such as business management and military, scenario planning in urban planning can be defined as a concept or tools that produce options through a combination of relevant resources and select the most feasible one to achieve the desired goals. As a single term, existing literature on scenario planning summarizes the characteristics of scenario planning. First, it is a "strategic" method to make flexible and predictable long-term plans. Second, it can make plans more "testable" at present through a set of evidence and values. Third, many quantifiable data and personal experiences or background of their ideal urban futures can serve as evidence and values in urban planning. Fourth, through scenario planning systems or tools, highly visual, descriptive, evaluative, prescriptive outcomes are shared with various planning parties and individuals.

During this period, adopting the concept and framework of scenario planning to the urban planning field and developing its tools are closely related to the development of indices and values in urban sociology and the potential of urban planning as disciplinary convergence in academia (Guhathakruta, 1999). Urban sociology emerged as one of the urban planning research fields in the 1960s, but due to its qualitative characteristics, it was not easy to develop some indices or values representing specific aspects of social phenomena. However, as planning research methods developed over time, planners produced some quantifiable social indices from the mid-1980s. Also, from the 1990s, planning literature about identifying a relationship between physical dimensions and non-physical variables (such as variables and data from public health) and developing integrated models were published in planning journals to estimate more comprehensive and accurate results of the future impacts of possible urban changes.

The main difference between scenario planning and former support systems (DSS and PSS) is that outputs generated from scenario planning tools can represent one comprehensive plan with a story behind it. This means that once planners collect data and input them, the performance of each scenario can be automatically calculated based on sets of models inside the system. The data management and processing in scenario planning tools are made through interaction between input data and outcomes, showing various impacts of each scenario (Brail and Klosterman, 2001; Brail, 2008). In scenario planning, data and information are not necessarily from statistical data resources developed by planners or organizations.

Scenario planning contains many assumptions that users need to create, and these input data can be generated through collaboration among users and observation of the current physical, environmental, economic, and even cultural conditions. Unlike former supporting systems, the presence of these scenario planning assumptions supports participatory planning and allows people to discuss what they expect to see in the future within the system framework. Through these data processing, personal background, knowledge, experience, and information can be used to prove the validity of their needs and goals (Schwartz, 1991; Ringland, 1998; Schoemaker, 1995; Ogilvy and Schwartz, 1998; Krizek and Forsyth, 2009).

Currently, there are four leading scenario planning tools in the U.S., and most of them (except for CommunityVIZ) are open-source scenario planning tools so that anyone can get the software and use them (Holway, Gabbe, Hebbert, Lally, Mathews, and Quay, 2012; Condon, Cavens, and Miller, 2009). Condon, Cavens, and Miller (2009) argue that characteristics of scenario planning tools can be identified based on six categories—scope, methodology, scale, and support for policymaking. In terms of scope, how many sectors scenario planning tools can use to construct and measure the future scenarios is critical. Some scenario planning tools currently used in the U.S. deal with only one sector to provide important quantitative baseline information on the specific aspect of the future scenarios. Leading scenario planning tools identified in this study all have multi-sector scope. In terms of methodology, scenario planning tools can be organized according to many approaches: spatial/non-spatial, top-down/bottom-up, simulation/end-state, and observation-based/process-based. Scale suggests at which geographic scales they can be operated and used. Finally, support for policymaking means to what extent of policy cycle model—five stages of decision-making process such as information gathering, interpretation, collaborative design, and policy formulation, implementation, and monitoring and evaluation—scenario planning tools can be covered. Through case studies of applications of three selected scenario tools in their planning projects—INDEX, Envision Tomorrow, and A Development Pattern Approach (DPA), the authors argued that future scenario planning tools should be evolved towards providing three-dimensional, multi-scale, policy-supportive, and iterative predictive models embedded within one integrated planning support system in an accessible and affordable way for all people affected by various predicted scenario planning outcomes (Condon, Cavens, and Miller 2009).

More recently, since 2010, the term GeoDesign is emerging among planners and geographic professionals. The concept and framework of GeoDesign were first officially introduced at the 2010 GeoDesign Summit. As the term represents, the basic concept of GeoDesign aims to integrate geo-information analysis and the design process. Therefore, as with planning support systems and scenario planning tools, the system structure of GeoDesign consists of databases from various external sources, the user interface system that allows planners and people to produce planning alternatives and sets of models that evaluate the performance or impacts of each alternative on the future urban development or growth. Also, users can reflect their design goals and assumptions into the system to produce options.

However, GeoDesign also has its distinctive features. First, it emphasizes the integration of the design process to support creativity and the image of the future environment. The main reason for the emergence of the GeoDesign concept is to improve flexibility in designing places rather than producing analytic outcomes. Scenario planning and PSS were effective in building up comprehensive sets of data and analytic models for planning alternatives, but the integration of design creativity and visualizing the future built environment as it looks were still a dilemma in these systems. Conceptually, GeoDesign aims to implement the design process within the system framework and reflect design values into traditional modeling and data processes. This is not what scenario planning tools are capable of, and GeoDesign attempts to fill this gap. Second, unlike PSS and scenario planning tools, models used in GeoDesign are designed based on the actual design process rather than models based on elements of the built environment. To evaluate the performance and impact of alternatives, GeoDesign consists of six process-based models—Representation Models, Process Models, Evaluation Models, Change Models, Impact Models, and Decision Models. As with scenario planning tools, the modeling process in GeoDesign also has inclusive, iterative, and scientific characteristics, which means users can intervene in the process any time, exchange feedback through review of that procedure and discussion, and conduct that procedure again or going back to the former procedures based on their feedback or planning goals (Steinitz, 2012). Third, from the system development perspective, GeoDesign is more concentrated on real-time interaction between users and the system. Although scenario planning can use big data to produce scenarios in a comprehensive and accurate way, the absence of standards for data development and management is one of the major challenges in scenario planning. To overcome this problem, GeoDesign systematically supports standardized databases. All databases relevant to GeoDesign are categorized and organized according to data management guidelines set up at the beginning of GeoDesign development and will be stored via the computer clouding system. When databases need updates, users can access this clouding system and change the data (Steinitz, 2012; McElvaney, 2012; ESRI, 2010).

As the most recent tools and concepts, more challenges will emerge in applying scenario planning and GeoDesign. To date, most of the challenges identified in existing literature are about how to implement them effectively in planning practice with various spatial levels and how to maintain databases

(Holway, Gabbe, Hebbert, Lally, Mathews, and Quay, 2012; Condon, Cavens, and Miller, 2009; Steinitz, 2012; McElvaney, 2012; ESRI, 2010).

FINDINGS AND DISCUSSION

Based on what this paper has reviewed so far, this paper concludes the evolution of data-driven tools in urban planning as follows. First, data-driven tools have been directed towards comprehensiveness. When scientific planning methods and approaches emerged in the early 1900s, planners thought that one single mathematical equation could give them one good solution to address the urban problem. Also, as a part of pure science and engineering, problems and issues in urban planning were dealt with structured but restricted mathematic equations. Aggregation of these outputs was considered as a comprehensive solution of urban plans. However, as time went by, the complexity and uncertain nature of a city and urban planning made planners reconsider pure scientific approaches and methods, and planners have devised data-driven tools that include more variables and analytic models over time. Although there was a time when the comprehensiveness of data-driven tools was seriously criticized due to the lack of technologies, data-driven planning tools have been designed so that planners and even people can set up goals and assumptions, input data, produce options, and evaluate performance and future possible impact of alternatives.

Second, problems and issues in data-driven tools have become more complicated and inclusive. Problems in data-driven tools until the 1970s include only one variable relevant to the expected output: for example, how will population change in cities? How does the number of workers change by distance from CBD? Also, to get a correct result based on a mathematical equation, most unscientific but influential factors or variables to the outputs were abandoned because of the inability of calculation. The application of outputs drawn from data-driven models during this period was seldom difficult, and there was a big gap between experimental outcomes and actual planning practice. The emergence of the decision support system (DSS) in the 1980s provided planners with opportunities to solve unstructured problems and support tools since then became capable of including unscientific and qualitative variables within the systems and developed integrated models that allowed planners to identify the relationship between scientific dimensions and non-scientific dimensions. As we saw in scenario planning and GeoDesign, these qualitative variables and data made possible planning alternatives more realistic and feasible.

Third, data-driven tools in urban planning have been developed by overcoming the limitations of former tools. Existing literature on scientific planning mentions that the emergence of scientific planning was in response to the City Beautiful Movement, which saw planning as an art (Nolen, 1909; Ford, 1913). Since then, data-driven tools in each time period have been designed to address limits in the former tools. For example, decision support systems were developed to produce outputs for poorly structured planning problems that neither scientific planning nor urban models covered. When scenario planning tools emerged, planners and people had a strong demand that tools include both scientific and non-scientific variables within the system framework through the development of integrated models.

Fourth, in terms of spatial scope, data-driven tools have expanded their applicability to various spatial and temporal levels. Except for the initial decision support system, most of the data-driven tools were devised by planners to estimate the possible future impact of planning options in large-scale areas such as cities, metropolitan areas, and regions. However, due to lack of data availability and analytic methods, impact assessment of large-scale areas using tools was realized when PSS tools were applied. Existing literature illustrates that GeoDesign can theoretically produce multiple design options up to the global level. Also, in terms of temporal level, advances in data-driven planning tools allowed planners to produce long-term planning alternatives. For example, most scenario planning tools used by planners and people can create scenarios for 30 or 50 years, and GeoDesign can produce up to 100-year planning options.

Finally, in terms of operating tools, data-driven tools have been developed towards including various planning parties and combining data from them together with traditional databases insider the systems. While data in traditional data-driven tools such as scientific planning and urban models were shared only by planners or organizations involved in the planning projects, advances in the capacity of personal computers enabled people to have access to these databases.

Particularly, big data and advanced data science techniques—such as machine learning, deep learning, artificial intelligence, and so on—can help planners examine a wide range of future urban growth alternatives and compare their performance or potential challenges effectively. Data-driven tools such as scenario planning and GeoDesign encourage the user discussion and review process by asking them to input goals and assumptions for evaluating performance and the impacts of alternatives they create later in the process. Table 1 summarizes a more detailed comparison of each data-driven tool identified in urban planning.

Table 1
A Comparison Table of Data-driven Tools in Urban Planning

Element	Scientific Planning	Urban Models	Decision Support Systems (DSS)	Planning Support Systems (PSS)	Scenario Planning (Tools)	GeoDesign
Year when the term emerged	Adopted to urban planning between the early 1900s and 1950s	Began to emerge in the late 1950 and widely used in the 1960s and early 70s	Began to widely used between the late 1970s and the 1980s	Widely used in the 1990s	Began to emerge in the late 1990s and widely used since the 2000s	The term officially emerged on Jan, 2010 in GeoDesign Summit.
Spatial Scope	Small-scale areas (Neighborhoods, CBD, etc.)	Large-scale areas (City, Metropolitan areas, Region)	Returning back to small-scale for short-term solutions	Large-scale areas (City, Metropolitan areas, Region)	Large-scale areas (City, Metropolitan areas, Region)	Large-scale areas (City, Metropolitan areas, Region)
Definition	One part of scientific fields that suggests urban problems can be solved through one or a set of mathematical equations.	Models that comprehend the basic activities of the urban area or those which deal with only a single major part of the scenario.	Models or tools that are designed to deal with poorly structured decisions by facilitating iterative and participative decision process.	A system with a wide diversity of geo-information dedicated to support planning processes at any particular spatial scale and within a specific planning context.	The process of producing options through a combination of resources and selecting the best one to achieve the desired goals defined by scenario planners.	A conceptual, invented word combining the terms “Geography” and “Design” The process of changing geography by design.

Table 1
A Comparison Table of Data-driven Tools in Urban Planning (continued)

Element	Scientific Planning	Urban Models	Decision Support Systems (DSS)	Planning Support Systems (PSS)	Scenario Planning (Tools)	GeoDesign
Features	It basically contains iterative and quite strong feedback process like a scientific experiment to build a theory. To modify the output, planners only correct one mathematical equations.	It is useful explaining one single planning phenomenon or figuring out structured planning problems that can prove validity through one set of math equations.	It is designed to analyze short-term policy making by isolated individuals and organizations.	With integration of more physical, economic, environmental, fiscal datasets and spatial information, more accurate outputs can be produced and visualized.	It allows users to computerize the future changes in the built environment based on their desired goals and assumptions in both mathematical and visual way.	This system can analyze current urban phenomena by analyzing big data and present the possible future changes in a highly detailed and visualized way on a real-time basis.
Dataset(s) Used	Highly descriptive data were used (ex. the number of population and workers)	Descriptive with datasets for the physical properties of the built environment added	Began to use geo-information datasets	Using integrating values or datasets drawn from outputs of models within the system	Began to use non-scientific and non-physical datasets from sociology.	Three-dimensional, real-time basis spatial information is emphasized for visualization.

Table 1
A Comparison Table of Data-driven Tools in Urban Planning (continued)

Element	Scientific Planning	Urban Models	Decision Support Systems (DSS)	Planning Support Systems (PSS)	Scenario Planning (Tools)	GeoDesign
Data Processing	Data gathering was very hard and conducted through field observation or survey.	Additional variables and equations were used for identifying their relationships in a specific problems.	Multiple criteria for evaluating outputs were applied in the data processing process.	Once input data enter, outputs are automatically calculated through a set of models inside the system.	By using formatted spreadsheets, outputs are automatically calculated through models inside the system.	Input data go through a set of process-based models. Each model can produce its own outputs.
Outputs	One simple, structured output (in numbers or tables)	Simple but more interrelated outputs from math equations.	Analytic results of a specific planning problem (in indices or values)	Analytic results of impacts of each alternatives with analytic maps	Visualized scenario maps and scenario performance outputs	two- or three-dimensional mapping and performance outputs

Table 1
A Comparison Table of Data-driven Tools in Urban Planning (continued)

Element	Scientific Planning	Urban Models	Decision Support Systems (DSS)	Planning Support Systems (PSS)	Scenario Planning (Tools)	GeoDesign
Limitations	It cannot explain "unscientific" variables and values only by using mathematical equations.	Generally useful in finding out solutions to "structured problems" (does not effectively explain the future for "poorly-structured" planning issues).	It does not effectively analyze long-term planning issues or problems by using this system model.	It succeeds in development of integrated models for scientific (physical/ economical/enviro-mental variables and data) and non-scientific (social values and variables). However, it still failed to include these models inside the system.	Data availability is a big issue. Not yet shared and spread out among individuals, organizations, and planners Additional trainings required due to the system complexity	For now, it only has a concept and a theoretical framework for development. Collecting and managing big data at both macro and micro spatial levels and integrating them into reality are very difficult to realize for now.

Combining all information communication technologies (ICTs) and data science techniques together, the concepts of the term "smart city digital twins (SCDT)" are recently receiving attention from planning practitioners and scholars. A smart city digital twin is defined as "a living digital replica of a city" whose changes are continuously updated using real-time big data analyses to get a holistic view of future urban changes and interactions among various aspects of cities (American Planning Association 2021). SCDTs aims to mirror the current conditions and aspects of a city itself as much as possible. Big data collected on a real-time basis are used to develop more comprehensive predictive models for enhancing decision-making and strengthening the city's capacity to adapt to a wide range of changes. In the SCDT approach, all different types of data—quantitative and qualitative data/structured and unstructured data—can be unified into one holistic data analytics and visualization environment, and all stakeholders can participate in informed decision-making processes by contributing themselves to data collection and analysis (Francisco et al. 2019; Mohammadi & Taylor 2017; Mohammadi et al. 2020). Through collaboration with the computer science field, there is an increasing number of planning journal papers being published using the advanced human- and infrastructure-generated data—such as crowdsourcing mobile data applications—and some Asian countries such as Singapore and South Korea (especially Seoul metropolitan city) are currently conducting pilot tests of their SCDT platform as a planning decision-making tool for testing a range of planning scenarios.

CONCLUSION

Conducting a literature review of data-driven tools in urban planning in the U.S., this paper aims to look at features of data-driven tools in urban planning in terms of forms, structures, functions, data management, and operation. This investigation aims at one question—how data-driven tools in urban planning have evolved. A wide-ranging literature review shows that there have been five significant transition phases and one major shift mainly caused by advances in computer and data processing technologies. Also, as shown in Table 1, this paper argues that each data-driven tool term has its dominant period, but these terms' basic concepts and principles have overlapped with one another. This also indicates that technological advances are significantly related to the development of data-driven tools throughout the urban planning history and changes the form, structure, and functionality of tools and their operation to generate more comprehensive and accurate outputs.

The nature of planning lies around the complexity and uncertainty of the future. To understand what we expect to see in the future, data-driven tools have played a role in producing a range of possible futures and evaluating their performance according to the goals and assumptions we define. As people's needs and demand for quality of life become diverse, more future challenges will lie ahead for planners, and planning tools or models will be constantly improved. However, even if data-driven tools are updated and can analyze the future in a more comprehensive and accurate way than before, they cannot predict exactly what will happen in the future. What is important here is that data-driven tools can bring about more efficient interaction among various planning parties to reach the consensus of the possible future.

REFERENCES

Abbott, J., (2005). Understanding and Managing the Unknown – The Nature of Uncertainty in Planning. Journal of Planning Education and Research, 24, pp. 237-251.

Adler, S. (1987). The New Information Technology and the Structure of Planning Practice. Journal of Planning Education and Research, 6(2), 93–98. doi:10.1177/0739456X8700600204

American Planning Association (2021). PAS Quick Notes: Smart Digital City Twins. PAS QuickNotes No. 89. Retrived on August 17, 2021 from https://www.planning.org/publications/document/9209455/.

Batty, M. (1994). A Chronicle of Scientific Planning: The Anglo-American Modeling Experience. Journal of the American Planning Association, 60(1), 7–16. doi:10.1080/01944369408975546

Batty M. (1995). Planning Support Systems and the New Logic of Computation, *Regional Development Dialogue*, 16(1), pp. 1-17.

Bloom M., Menefee M. (1994). Scenario Planning and Contingency Planning, *Public Productivity & Management Review*, Vol. 17, no. 3, pp. 223- 230.

Bradfield, R., Wright, G., Burt, G., Cairns, G., Ven Der Heijden, K. (2005). The origins and evolutions of scenario techniques in long range business planning, *Futures*, 37, pp. 795-812.

Bonczek, R. H., Holsapple, C. W., & Whinston, A. B. (1979). Computer-Based Support of Organizational Decision Making*. Decision Sciences, 10(2), 268–291. doi:10.1111/j.1540-5915.1979.tb00024.x

Camarinhas, C. T. F. (2011). The Construction of Modern Scientific Urban Planning: Lisbon under French Urbanisme Influence (1904-1967). Planning Theory & Practice, 12(1), 11–31. doi:10.1080/14649357.2011.546667

Casseres, J. M. de. (1937). Principles of Planology: A Contribution to the Scientific Foundation of Town and Country Planning. The Town Planning Review, 17(2), 103–114.

Chermack, T. J., Lynham, S. A., Ruona, E. A. (2001). A review of scenario planning literature, *Futures Research Quarterly*, 7(2), pp. 7-32.

Christensen K. (1985). Coping with Uncertainty in Planning. Journal of the American Planning Association, Winter 1985, pp. 63-71.

Condon P., Cavens D., and Miller N. (2009). Urban Planning Tools for Climate Change Mitigation. A Policy Focus Report of Lincoln Institute of Land Policy.

Connor, D. S., & Storper, M. (2020). The changing geography of social mobility in the United States. *Proceedings of the National Academy of Sciences*, *117*(48), 30309-30317.

Fairfield, J. D. (1994). The Scientific Management of Urban Space Professional City Planning and the Legacy of Progressive Reform. Journal of Urban History, 20(2), 179–204. doi:10.1177/009614429402000202

Francisco, A., Mohammadi, N., & Taylor, J. E. (2019). Smart City Digital TwinEnabled Energy Management: Toward Real Time Urban Building Energy Benchmarking. *Journal of Management in Engineering, Special Collection on Engineering Smarter Cities with Smart City Digital Twins*, 36(2): 04019045.

Franklin, D. (2004). The Basics of Scenario-based Forecasting and Planning. Strategic Analysis, Vol. 4, No. 1, pp. 1-6.

Forsyth, A. (2012). Commentary: Alternative Cultures in Planning Research—From Extending Scientific Frontiers to Exploring Enduring Questions. Journal of Planning Education and Research, 32(2), 160–168. doi:10.1177/0739456X12442217

Frego Associates. (2013). Envision Tomorrow Plus: Scenario Indicator + Apps. Available from http://www.frego.com/etwiki/doku.php?id=support

Global Business Network. (2002). *Developing and Using Scenarios.* (San Francisco, Global Business Network).

Godet, M. and Roubelat, F. (1996). Creating the future: the use and misuse of scenarios, *Long Range Planning*, 29(2), pp. 164-171.

Godet, M. and Roubelat, F. (2000). Scenario planning: an open future, *Technological Forecasting and Social Change*, 65, pp. 1-2.

Green, T. L. (2015). Places of inequality, places of possibility: Mapping “opportunity in geography” across urban school-communities. *The Urban Review*, 47(4), 717-741.

Guhathakurta, S. (1999). Urban Modeling and Contemporary Planning Theory: Is There a Common Ground? Journal of Planning Education and Research, 18(4), 281–292. doi:10.1177/0739456X9901800401

Hackathorn, R. D., & Keen, P. G. W. (1981). Organizational Strategies for Personal Computing in Decision Support Systems. MIS Quarterly, 5(3), 21–27. doi:10.2307/249288

Haigh, T. (2011). The history of information technology. Annual Review of Information Science and Technology, 45(1), 431–487. doi:10.1002/aris.2011.1440450116

Han, S.-Y., & Kim, T. J. (1989). Can Expert Systems Help with Planning? Journal of the American Planning Association, 55(3), 296–308. doi:10.1080/01944368908975417

Harris, B. (1989). Beyond Geographic Information Systems. Journal of the American Planning Association, 55(1), 85–90. doi:10.1080/01944368908975408

Harris, B. (1994). The Real Issues Concerning Lee’s “Requiem.” Journal of the American Planning Association, 60(1), 31–34. doi:10.1080/01944369408975548

Heiman, M. K. (1997). Science by the People: Grassroots Environmental Monitoring and the Debate Over Scientific Expertise. Journal of Planning Education and Research, 16(4), 291–299. doi:10.1177/0739456X9701600405

Holway J., Gabbe C., Hebbert F., Lally J., Matthews R., and Quay R. (2012). Opening Access to Scenario Planning Tools. A Policy Focus Report of Lincoln Institute of Land Policy.

Iacono, M., Levinson, D., & El-Geneidy, A. (2008). Models of Transportation and Land Use Change: A Guide to the Territory. Journal of Planning Literature, 22(4), 323–340. doi:10.1177/0885412207314010

Innes, Judith E. (2004). Consensus building: clarifications for the critics, *Planning Theory*, pp. 5-20.

Inndes, Judith E. (1998). Information in communicative planning, *Journal of the American Planning Association*, pp. 52-63.

Kim, T. J. (2013). Scientific Reasoning and Methods in Urban Planning. International Regional Science Review, 36(1), 36–43. doi:10.1177/0160017612441365

Klosterman, R E, & Landis, J. D. (1988). Microcomputers in U.S. planning: past, present, and future. Environment and Planning B: Planning and Design, 15(3), 355–367.

Klosterman, Richard E. (1992). Evolving Views of Computer-Aided Planning. Journal of Planning Literature, 6(3), 249–260. doi:10.1177/088541229200600303

Klosterman, Richard E. (1994a). An Introduction to the Literature on Large-Scale Urban Models. Journal of the American Planning Association, 60(1), 41–44. doi:10.1080/01944369408975550

Klosterman, Richard E. (1994b). Large-Scale Urban Models Retrospect and Prospect. Journal of the American Planning Association, 60(1), 3–6. doi:10.1080/01944369408975545

Krizek, K., Forsyth A., Slotterback, C.A. (2009). Is there a role for evidence-based practice in urban planning and policy?, *Planning Theory & Practice*, 10(4), pp. 459-478.

Laplante, M. (2010). Evidence-based urban planning. Platnetizen, Available from http://www.planetizen.com/node/46699 (Retrieved September 14, 2012).

Lee, D. B. (1973). Requiem for Large-Scale Models. Journal of the American Institute of Planners, 39(3), 163–178. doi:10.1080/01944367308977851

Lee, D. B. (1994). Retrospective on Large-Scale Urban Models. Journal of the American Planning Association, 60(1), 35–40. doi:10.1080/01944369408975549

Madanipour, A. (2010). The limits of scientific planning: Doxiadis and the Tehran Action Plan. Planning Perspectives, 25(4), 485–504. doi:10.1080/02665433.2010.505066

Margo, R. A. (1986). Race, educational attainment, and the 1940 census. *The Journal of Economic History*, *46*(1), 189-198.

Mazzotti, F. J., & Morgenstern, C. S. (1997). A scientific framework for managing urban natural areas. Landscape and Urban Planning, 38(3–4), 171–181. doi:10.1016/S0169-2046(97)00032-7

Mohammadi, N., Vimal, A., & Taylor, J. E. (2020). Knowledge Discovery in Smart City Digital Twins. *In Proceedings of the 53rd Hawaii International Conference on System Sciences –* HICSS 53, Maui, HI.

Mohammadi, N., and Taylor, J. E. (2017). Smart City Digital Twins. *In Proceedings of the 2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, Honolulu, HI

Ogilvy J. and Schwartz P. Plotting your Scenarios (2008), An Exerpt from Learning from the Future by Fahey L., Randall. pp. 1- 18.

Pack, J. R. (1975). The Use of Urban Models: Report on a Survey of Planning Organizations. Journal of the American Institute of Planners, 41(3), 191–199. doi:10.1080/01944367508977877

Platt, S., Cooper, I. (2006). "Democracy and development" in *Building on the Past*, The Policy Press, pp. 239-261.

Preece, R. (1990). Development control studies: Scientific method and policy analysis. The Town Planning Review, 61(1), 59–74.

Rautenstrauch, W. (1945). What Is Scientific Planning? Philosophy of Science, 12(1), 8–18.

Sagasti, F. R. (1973). Towards a new approach for scientific and technological planning. Social Science Information, 12(2), 67–95. doi:10.1177/053901847301200204

Salingaros, N. A. (1998). A scientific basis for creating architectural forms. Journal of Architectural & Planning Research, 15(4), 283–293.

Sanoff, H. (2000). *Community Participation Methods in Design and Planning* (New York, John Wiley & Sons, Inc.).

Schoemaker, Paul J.H.(1995). Scenario planning: a tool for strategic thinking, *Sloan Management Review*, 36(2), pp. 25-40.

Schwartz, P. (1996). *The Art of The Long View*. (New York, Doubleplay).

Schwenk, C.R. (1988). *The Essence of Strategic Decision Making*, (Toronto, Lexington Books).

Shepard, G. H. (1937). SCIENTIFIC PLANNING IN CITIES. Sigma Xi Quarterly, 25(1), 10–18.

Silva, E., & Wu, N. (2012). Surveying Models in Urban Land Studies. Journal of Planning Literature, 27(2), 139–152. doi:10.1177/0885412211430477

Solesbury, W. (2002). The ascendancy of evidence, *Planning Theory & Practice*, 3(1). pp. 89-104.

Sprague, R. H., Jr. (1980). A Framework for the Development of Decision Support Systems. MIS Quarterly, 4(4), 1–26. doi:10.2307/248957

Taylor, N. (1999). Anglo-American town planning theory since 1945: three significant developments but no paradigm shifts. Planning Perspectives, 14(4), 327–345. doi:10.1080/026654399364166

Topalov, C. (1990). Scientific Urban Planning and the Ordering of Daily Life The First "War Housing" Experiment in the United States, 1917-1919. Journal of Urban History, 17(1), 14–45. doi:10.1177/009614429001700103

Trochim, W. M. K. and Donnely, J.P. (2008). Research Methods Knowledge Base. (Mason, Cengage Learning).

Turner, R. H. (1954). Occupational patterns of inequality. *American Journal of Sociology*, *59*(5), 437-447.

Varum, C. A., & Melo, C. (2010). Directions in scenario planning literature – A review of the past decades. Futures, 42(4), 355–369. doi:10.1016/j.futures.2009.11.021

Wates, N. (2008). *The Community Planning Event Manual* (Earthscan, London).

Wates, N. (2000). *The Community Planning Handbook* (Earthscan, London).

Wilson, A. (2010). The general urban model: Retrospect and prospect. Papers in Regional Science, 89(1), 27–42. doi:10.1111/j.1435-5957.2010.00282.x

Wu, N. and E. A. Silva. (2012). Surveying Models in Urban Studies. *Journal of Planning Literature*, 27(2), pp. 139-152.

Wulf T., Meiβner P., Stubner S. (2010). A Scenario-based Approach to Strategic Planning – Integrating Planning and Process Perspective of Strategy, *HHL Working Paper No. 98 in Leipig Graduate School of Management*, pp. 1-35.

Yeh, A. G. O. and Shi, X. (1999). Applying case-based reasoning to urban planning: a new planning-support systems tool. *Environment and Planning B: Planning and Design*, 26, pp. 101-115.