Jul 12th, 3:10 PM - 3:30 PM

A Survey on Data-driven Decision Support Systems Using Effective Narrative Pattern

Martin Komenda
Institute of Biostatistics and Analyses, Faculty of Medicine, Masaryk University, komenda@iba.muni.cz

Matěj Karolyi
Research Centre for Toxic Compounds in the Environment, Masaryk University, karolyi@iba.muni.cz

International Environmental Modelling and Software Society (iEMSs)
8th International Congress on Environmental Modelling and Software
Toulouse, France, Sabine Sauvage, José-Miguel Sánchez-Pérez, Andrea Rizzoli (Eds.)
http://www.iemss.org/society/index.php/iemss-2016-proceedings

A Survey on Data-driven Decision Support Systems Using Effective Narrative Pattern

Martin Komenda*[1], Matěj Karolyi[1,2]
1 Institute of Biostatistics and Analyses, Faculty of Medicine, Masaryk University
2 Research Centre for Toxic Compounds in the Environment, Faculty of Science, Masaryk University
{komenda, karolyi}@iba.muni.cz

Abstract: Nowadays information technology assists with decision making in various fields of human activities. This paper focuses on a survey of decision support portals and systems that provide any data visualization form and reporting in the environmental research field. Generally, data outputs in environmental research are too large, and it might be difficult for the general public to be well oriented when exploring shown data sets. Therefore, it would be appropriate to guide readers, or to point out the fundamental features of the presentation. One of the effective techniques to lead the reader is data visualization storytelling. It expands exploratory data analysis by an explanatory part, which provides a space for a really strong impact by the author. We discuss the real benefits of design strategies for narrative visualization in relation to information systems supporting the decision-making process. A deep survey presents here a global summary on particular decision support systems in terms of using data-driven document approach together with various forms of narrative patterns.

Keywords*:* data-driven approach, narrative pattern, martini glass structure, visualization.

1 INTRODUCTION

The power of progressive modern information and communication technologies can play an important role for constructing tools embedding algorithms, graphical interfaces and access to remote data through the Internet, Santos et al. (2011). Over the last two decades, different kinds of information systems have been developed for different purposes, depending on the need of various domains of human interests. The idea of a new dimension of decision-making model describing a limitation of a human being's rationality was introduced for the first time by Herbert Simon in 1950s. The model suggested that when an individual makes a decision, he examines a limited set of possible alternatives rather than all available options, Simon(1957). The theoretical background should be always enhanced with a real implementation in practice. Properly designed computer-based systems are able to automatically compile useful information from a combination of raw data, documents, personal knowledge or business models, to identify and to solve problems. According to Giordano et al. (2011). The so-called Decision Support Systems (DSS) constitute a specific class of these systems supporting business and organizational decision-making activities. Using the mode of assistance as the criterion, a taxonomy created by Power(2002) differentiates the following DSS: (i) communication-driven systems supporting more than one person working on a shared task; (ii) data-driven systems emphasizing access to and manipulation of time series of internal company data, and sometimes external data; (iii) document-driven systems managing, retrieving, and manipulating unstructured information in a variety of electronic formats; (iv) knowledge-driven systems providing specialized problem-solving expertise stored as facts, rules, procedures, or in similar structures; (v) model-driven systems emphasizing access to and manipulation of a statistical, financial, optimization, or simulation model. We paid a close attention only to the category of data-driven DSS, where file drawer and management reporting systems, data warehousing and analytical systems including business intelligence solutions are involved. These systems usually use data that have been extracted from all relevant internal and external sources. Supporting decision-making means that computerized tools are used to make sense of available data. Involved stakeholders can analyse, display and manipulate large structured data sets containing numeric and short character strings, and allow them to identify facts and draw conclusions about relationships and trends.

The way in which information is delivered to end users through a particular DSS is considered to be a fundamental issue. The complexity of the environment, the time scale and the diversity of environmental effect is such that the implication of decisions affecting the environment are beyond the imagination of most people involved. Instead of trying to explain all the inputs and outputs of a problem to people engaged in a decision, an object (in the form of a learning device) can be made available to these people, enabling them to play around with a problem. By including the available knowledge of specific problems, these systems allow experimentation without fatal consequences to the environment, Janssen(2012). This paper aims to review existing data-driven systems for supporting decision-making in environmental fields, and to show the importance of the presentation form together with narrative pattern structures.

2 VISUALIZATION AS STORYTELLING

Visualization is usually perceived as the visual representation of available data in very heterogeneous form, i.e. information in a text, numerical or illustrated format. From the perspective of environmental research, modern technologies allow us to visualize any kind of interdisciplinary data that integrates physical, biological and information sciences to the study of the environment, and the solution of environmental problems. Today, the trend aims to use various text and data mining techniques to explore information-rich data from huge databases and warehouses and create online visualizations. All types of structured interpretation of information can be easily visualized (static and dynamic graphs, diagrams, maps, information illustrations, etc.). The need to visualize data has emerged from various research areas, and data visualization has been a useful tool to the study of problems across the scientific research. Understanding all pieces that compose a visualization is important in order to evaluate the progress of the maturing visualization field, to help focus and direct future research, and to help to create better visualizations that make use of the elements that are essential for a visualization to become successful, Figueiras(2013). Segel et al. (2010) introduced a framework suggesting design strategies for narrative visualization including promising under-explored approaches for telling stories with data graphics. Storytelling and visual expression are integral parts of human culture and are closely connected to the narrative visualization.

2.1 Narrative visualization

Narrative visualizations combine conventions of communicative and exploratory information visualization to convey an intended story. It is used for understanding how design techniques that prioritize particular interpretations in visualizations that "tell a story" can significantly affect end-user's interpretation, Hullman et al. (2011). It can be explained as a non-visual mechanism that assist clear data representation and facilitate proper decision-making processes. The design of narrative visualizations can influence the interpretation in diverse ways, such as by presenting a preliminary statistic before a decision or by manipulating the anchor points on a survey scale, Schwarz et al. (1985). According to Segel et al. (2010), there are three ways of distinguishing categories of narrative visualizations: (1) genres; (2) visual narrative tactics that direct attention, guide view transitions, and orientate the user; and (3) narrative structure tactics such as ordering, interactivity, and messaging.

Effectivity of telling the stories through data visualizations mostly depends on the chosen technique. Which type of data visualization should be used is an essential decision. For example, data can be displayed by various graphs, but not all of them are suitable in all cases. If a wrong way is chosen, the visualization becomes ineffective, and the probability of presenting misleading data and making bad decisions is rising, Knaflic(2015). Construction of an appropriate narrative is the next element in the process of information transfer from an author to a reader. We considered the most powerful methods to show readers the approach (explanatory part), and then let them explore the data sets by their own steps (exploratory part). The so-called martini glass structure is one narrative pattern which uses this concept.

2.2 Martini glass structure

This concept follows a tight narrative path early on (stem of the glass) and then opens up later for free exploration (body of the glass). Different features of the visualization ensure that the end user does not get lost in the narrative during this exploratory stage, with orientation provided by the consistency of the visual platform, the updating progress bar, and the timeline slider, Herman(2011). The martini glass has a three-stage structure (see Figure 1a): (1) the attention of the user is focused on the entire domain, which will be introduced during the narrative (kicker). (2) The user goes through a relatively heavily author-driven scenario, in which the visualization is introduced through the use of text, annotations, nicely crafted animations, or interesting and evocative views (explanatory part). (3) When the author's intended narrative is complete, the user is put in charge and can actively explore the visualization following whichever path the user considers most interesting (exploratory part), Boy et al. (2015). Before the first stage, there is is starting point called the kicker, from which the user is automatically guided to the author-driven part, Man(2011). Visual storytelling together with appropriate graphical techniques, interactivity and highlighting can enforce the narrative flow and produce varying balances of presented information based on data-driven experiences.

2.3 Data-driven documents

Nowadays there are many ways of obtaining data, cleaning them and processing them. These data can be connected with every aspect of human life (behaviour, health, attitudes, etc.), scientific research (biology, chemistry, car industry etc.) or environmental protection (indicators, reporting). Collected observations are essential for making final decisions and for creating plans for the future. This procedure is called data-driven decision making, and data-driven documents are its cornerstones. On the other side, this document should be an output of data-driven development (D3). The structure of document is unspecified: it could be plain text, source code or multimedia file. The fact that a major part of its content is data-oriented is more important. As a specific example, D3 is an approach to visualization of websites. It makes it possible to bind input data to document object model (DOM), and to apply dynamic transforms to create, to update and to remove content, Bostock et al. (2011). The document is present in the whole project – from the start of development to the end of decision process (see Figure 1b).

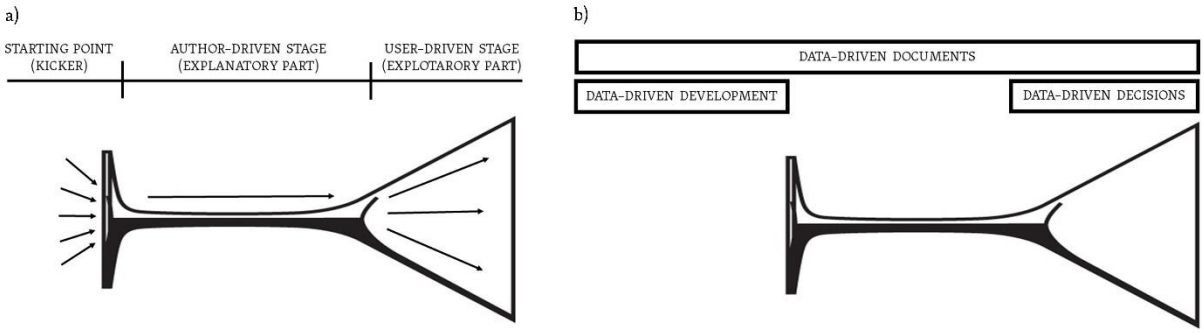

Figure 1. a) Structure of martini glass narrative, b) Martini glass in relation with data-driven approach.

4 CURRENT CHALLENGES

The promise of data-driven decision-making is now being recognized broadly, and there is growing enthusiasm for the effective way of delivering information to end users, Labrinidis et al. (2012). The understanding of data-driven decision-making strategies to make well-founded decisions on the basis of data is quite challenging. Scientists have realized that building expertise and capacity for data-driven decision-making is a necessary but not a sufficient condition for success, Wohlstetter et al. (2008). In our deep survey, we focus on stable environmental decision support systems on national or international level, which are developed and supervised by governmental organizations. These systems primarily serve as guaranteed information channels for the general public and may bring only proven and trustworthy data in a meaningful context. In order to comply with the established conventions of research, the following research question was stated: How often do environmental decision support systems combine data-driven approach with any form of narrative structure?

5 METHODS

The deployment of decision support systems is applied in practice in various fields of environmental research. In the following section, we will briefly introduce the most interesting web-based solutions using data-driven approach together with narrative structures falling into governmental, national or international domain.

The European Environment Agency (EEA) is an agency of the European Union, currently involving 33 member countries. Its system (available at http://www.eea.europa.eu) aims to provide sound, independent information on the environment. It provides a major information source for those involved in developing, adopting, implementing and evaluating environmental policy, and also the general public, Nelson(1999). EEA's mandate is: (i) to help the community and member countries make informed decisions about improving the environment, integrating environmental considerations into economic policies and moving towards sustainability; (ii) to coordinate the European environment information and observation network. Its range of provided data visualizations is wide: from static author-driven images through interactive filterable charts to even more complex narratives. Figure 2 shows an interactive map with automatic timeline filter. Unfortunately, the kicker stage is absent but this visualization is close to a martini glass narrative.

Figure 2. Interactive map including a filtering toolbox.

The National Climate Assessment summarizes the impacts of climate change on the United States, now and in the future. The system (available at http://nca2014.globalchange.gov) collects, integrates, and assesses observations and research from around the country, helping us to see what is actually happening and understand what it means for the general public. The interactive report includes analyses of impacts on seven sectors – human health, water, energy, transportation, agriculture, forests, and ecosystems – and the interactions among sectors at the national level. The website content is in the form of an interactive slideshow. The user is systematically guided through topics by author, and at some parts (see Figure 3) he/she is able to explore data in his/her own way.

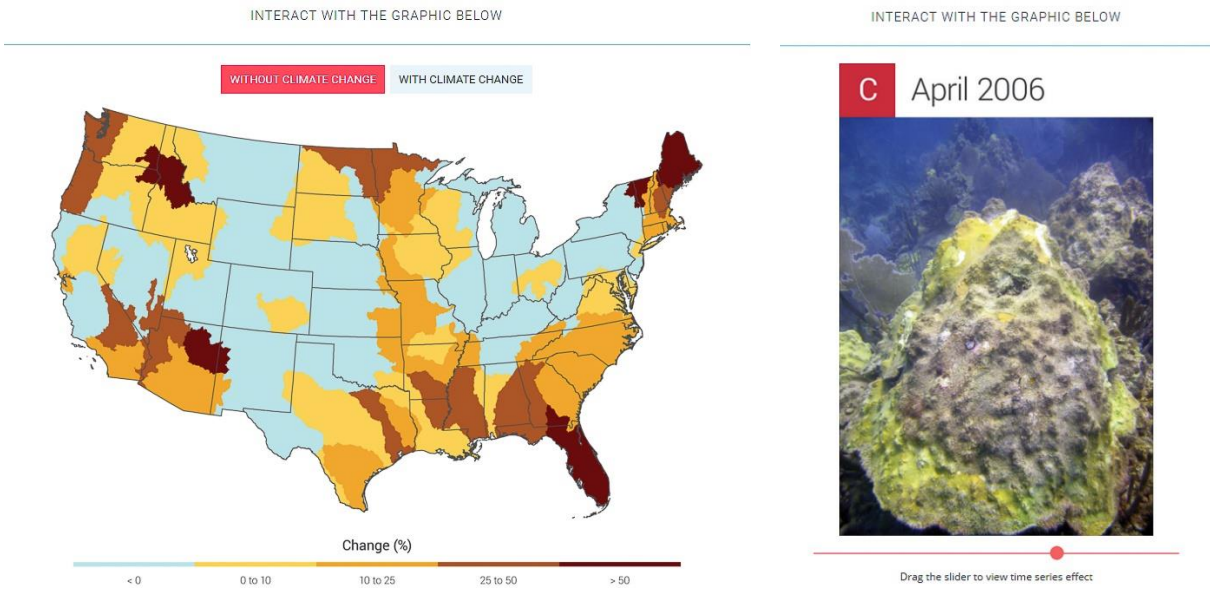

Figure 3. Difference view map and time changes visualization.

GENASIS (Global environmental assessment and information system) is a tool developed by expert teams of the Research Centre for Toxic Compounds in the Environment (RECETOX) and the Institute for Biostatistics and Analyses (IBA) of the Masaryk University. The aim of GENASIS system (available at http://www.genasis.cz/index-en.php) is to compile validated data on persistent organic pollutants, including their properties, sources, long-term levels, life times, transport mechanisms, effects and risks, scattered throughout various institutions and ministries, and to provide tools for their visualization, analyses, interpretation, assessment of environmental and human risks or modelling of fate, Holoubek et al. (2011). Data are presented in a form of map customizable by very detailed data selection (see Figure 4) or summary distribution graphs. Data selection depends completely on the user and requires his/her considerable knowledge in the field.

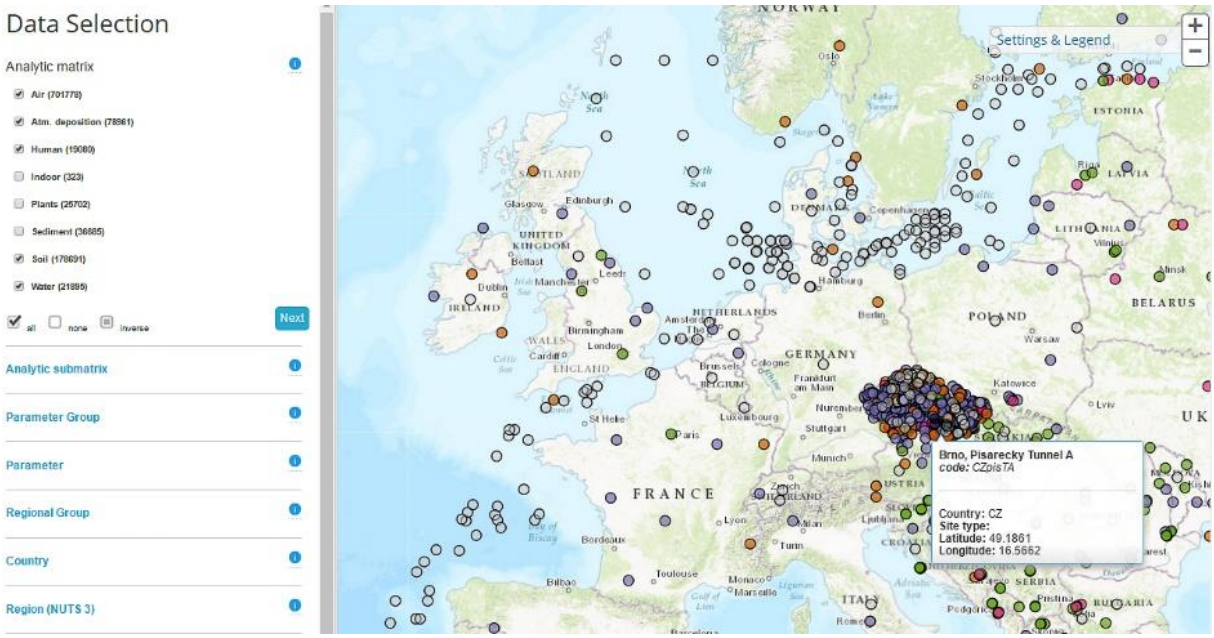

Figure 4. GENASIS Data browser.

Report on the Environment (ROE) by the Environmental Protection Agency (available at https://cfpub.epa.gov/roe) is an interactive resource showing how the condition of the environment and human health in the United States has been changing over time. Targeted for anyone interested in environmental trends, the ROE presents the best available indicators of national trends in air, water,

land, human exposure and health, and ecological condition, US EPA (2016). Presentation of data is mostly user-driven, but each visualization has a comprehensive and structured author's commentary, which has a unified form and gives readers a basis for further exploration.

6 SUMMARY ON DECISION MAKING SYSTEMS

After surveying issues in the paragraphs above, we present here a global summary on particular decision support systems in terms of using D3 approach together with various form of narrative (see Table 1). Basic narrative structures shown in the table are author-driven and user-driven. The martini glass and the interactive slideshow are a combination of these two approaches. Websites and systems where the report outputs are static downloadable documents (such as PDF files) were excluded from our survey. Visualizations in the remaining systems were divided into three categories: (1) a static web page – mostly images and texts with no interactions provided; (2) animated web page – visualization performs any type of automatic movement to improve readers focus on important data; and (3) interactive web page – reader is able to interact with charts, filter them, sort them, etc.

Table 1. Overview of decision support systems including selected narrative parameters.

Name	D3 support	Narrative structure	Visualization form	Description
Australia Government - Department of the Environment	Yes	User-driven	Interactive	A small section of the portal is dedicated to an interactive map tool showing regions of Australia.
Environmental Protection Agency	Yes	User-driven	Interactive	The portal provides various types of filterable graphs showing data about different environmental fields (air, water, land etc.) Textual information supporting graphs has a standardized structure and provides a strong knowledge background by the authors.
European Environment Agency	Yes	User-driven / Author driven (martini glass)	Interactive, Animated	The portal includes interactive user-driven visualizations; and some of them can be changed to an automatic mode, which unites author- and user-driven narrative to martini glass.
GENASIS Environmental Data Repository	Yes	User-driven	Interactive	Data browser is part of the GENASIS system. This feature shows various data-enhanced maps and other types of plots customizable by a set of settings and specified by data selection.
Global Environmental Facility	Yes	User-driven	Interactive	An interactive map showing projects across the world is extended by statistical pie charts. These graphs include statistics for actual filter settings, and are updated after every user change.
Intergovernmental Panel on Climate Change	No	Author-driven	Static Web Page	Reports presented mostly by publications (technical papers, supporting materials), but images of plot in jpg files are also present on the website. A good overall context about presented data is not available for the user.
International Waters Learning Exchange and Resource Network	Yes	User-driven	Interactive	A visualized output of results is shown on a customized Google Map. Project results can be filtered, and information on the map is usually expanded by a supporting graph, which is referenced in a tooltip.
Japan Meteorological Agency	Yes	User-driven	Static Web Page	The web page provides actual information about dangerous weather and earthquakes across Japan. Data are presented by images with support of several filters.
Ministry for the Environment	Yes	User-driven	Interactive	Very detailed reports over the map of New Zealand. Users can add graphical interpretation of stored data to the map and explore it their own way.

Name	D3 support	Narrative structure	Visualization form	Description
National Climate Assessment	Yes	User-driven / Author-driven (Interactive slideshow)	Interactive	The portal provides various types of interactive graphs. The user can apply filter tools to get subsets of presented data. Author-driven narrative is present in static graph images extended by a large amount of supporting textual comments.
Oceanic and Atmospheric Administration	Yes	User-driven	Interactive	This portal provides customizable visualizations, which are completely driven by the user. The set of filters is large, allowing to plot countless graph combinations above the provided data.
The New Climate Economy	No	Author-driven	Static Web Page	The author is mostly interacting with users through a video, mentioning keywords and data around a specific topic. After watching a video, the user can explore data visualizations through static image graphs. This web page has a significant potential to become a martini glass narrative structure.
World Climate Report	No	Author-driven	Static Web Page	Mostly textual explanation of the presented problem. Graphs are presented via static images.
World Resource Institute	Yes	Author-driven / User-driven	Interactive	The majority of visualizations are static images. Climate Data Explorer represents the interactive part of portal: the user can explore the map enhanced by filter tools.

7 CONCLUSIONS

This paper provides a survey describing the most interesting decision support systems used at national or international level in terms of various narrative paths. The introduced concept of narrative structure, together with data-driven approach can provide a robust methodological and technological background for an effective delivery of information, which are both required to design decision making support systems. The shown concept can be quite helpful for researchers working in the field of environmental informatics in situations where a large amount of data must be presented in the most comprehensible way. The summarized overview shows the real implementation of selected web-oriented solutions in environmental domains. In some cases, developers have properly used narrative structural framework as a strategy for presenting the content of their story to a reader, a listener or a viewer. In general, narratives are an integral part of human expression, Cohn(2013). From the audience's perspective, attractive graphical representation (interactive and animated interpretation of data) can offer a better understanding. Data-driven approach, using effective narrative patterns in combination with dynamic visualization components, provide the future for environmental decision support systems as well as all domains, where the comprehensibility of presented information is required.

REFERENCES

Bostock, M., Ogievetsky, V., and Heer, J. 2011. D3 data-driven documents. Vis Comput Graph IEEE Trans On. 17:2301–2309.

Boy, J., Detienne, F., and Fekete, J.-D. 2015. Storytelling in Information Visualizations: Does it Engage Users to Explore Data? In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. ACM. pp. 1449–1458.

Cohn, Neil. 2013. Visual narrative structure. Cognitive science. pp 413-452.

Figueiras, A. 2013. A Typology for Data Visualization on the Web. In: 2014 18th International Conference on Information Visualisation. Los Alamitos, CA, USA: IEEE Computer Society. pp. 351–358.

Giordano, J.O., Fricke, P.M., Wiltbank, M.C., and Cabrera, V.E. 2011. An economic decision-making support system for selection of reproductive management programs on dairy farms. J Dairy Sci. 94:6216–6232.

Herman, D. 2011. Basic elements of narrative. John Wiley & Sons.

Holoubek, I., Dusek, L., Klánová, J., Kubásek, M., Jarkovskỳ, J., Baros, R., Komprdová, K., Bednárová, Z., Hulek, R., and Hrebícek, J. 2011. GENASIS Information System: A Global Environmental Assessment of Persistent Organic Pollutants. In: ISESS. Springer. pp. 480–485.

Hullman, J., and Diakopoulos, N. 2011. Visualization Rhetoric: Framing Effects in Narrative Visualization. IEEE Trans Vis Comput Graph. 17:2231–2240.

Janssen, R. 2012. Multiobjective decision support for environmental management. Springer Science & Business Media.

Knaflic, C.N. 2015. Storytelling with Data: A Data Visualization Guide for Business Professionals. John Wiley & Sons.

Labrinidis, A., and Jagadish, H.V. 2012. Challenges and opportunities with big data. Proc VLDB Endow. 5:2032–2033.

Man, P. 2011. Narrative Structures in Data Visualizations to Improve Storytelling | Masters of Media.

Nelson, D.A. 1999. European Environment Agency. Colo J Int Environ Law Policy. 10:153.

Power, D.J. 2002. Decision Support Systems: Concepts and Resources for Managers. Greenwood Publishing Group.

Santos, L., Coutinho-Rodrigues, J., and Antunes, C.H. 2011. A web spatial decision support system for vehicle routing using Google Maps. Decis Support Syst. 51:1–9.

Schwarz, N., Hippler, H.-J., Deutsch, B., and Strack, F. 1985. Response scales: Effects of category range on reported behavior and comparative judgments. Public Opin Q. 49:388–395.

Segel, E., and Heer, J. 2010. Narrative visualization: Telling stories with data. Vis Comput Graph IEEE Trans On. 16:1139–1148.

Simon, H.A. 1957. Administrative Behavior: A study of decision-making processes in adminstrative organization.

US EPA, O. of R. and D. 2016. EPA's Report on the Environment.

Wohlstetter, P., Datnow, A., and Park, V. 2008. Creating a system for data-driven decision-making: applying the principal-agent framework. Sch Eff Sch Improv. 19:239–259.