(Un)intended Value Implications of Graphical Representations of Data

Milena Radzikowska, MDes, PhD, Mount Royal University, Calgary, Alberta, mradzikowska@mtroyal.ca

Stan Ruecker, MDes, PhD, University of Illinois, School of Art and Design, Champaign, IL, USA, sruecker@illinois.edu

Abstract

The design of meaningful graphical objects to represent collection items must balance the following: amount of useful information that can be communicated through the object's graphical form, meaningful graphical difference between individual items or groups of items, and restraint in form complexity to allow for the simultaneous display of numerous collection items at a small size. How the user interprets difference and sameness and, more importantly, whether the user attaches hierarchical value to the emergent categories, may play a significant role in determining whether that user focusses attention on one set of data over another, on one set of processes over another, and ultimately, on one set of tasks over another. This paper examines the significant consequences for the understanding of the user resulting from representation of data, files, and other objects in a human-computer interface (HCI), and proposes that new approaches may be indicated, given the growing complexity of what is being represented and how what is represented can be used.

Keywords: *visualization, HCI, design*

Introduction

In his 1986 classic, the *Whale and the Reactor,* Langdon Winner argues that it is important to interrogate the political implications of design. He points out, for example, that the low clearance of bridges, over parkways in Long Island, New York, were a deliberate choice made by city planner, Robert Moses, to make the park inaccessible to lower-income groups, who relied on public busses for transportation. In a similar vein, Victor Papanek in his 1973 edition of *Design for the Real World,* calls design in the age of mass production "the most powerful tool with which man shapes his tools and environment (and by extension, society and himself)" (14), then accuses it of putting "murder on the level of mass production" (14). He states that, while advertising designers persuade "people to buy things they don't need, with money they don't have", industrial designers create unsafe, unnecessary, "tawdry idiocies" to be "hawked by advertisers" (14). As evidence he points to the industrial process and product-use that create exorbitant waste material, pollute our air and water, and are capable of causing injury and harm to a cross-global population.

Though most human-computer interfaces (HCIs) are not the outcome nor the mass producer of industrial design, many enable mass production, distribution, purchasing, and obsolescence on a scale that does not have its equal in a physical counterpart. Take amazon.com as an example. In 2014, Amazon reported almost US$89 billion in net sales, with almost 114,000 total office and warehouse units, 181.12 million unique monthly visitors, and 305,258,547 unique products (Statista, 2015). While Amazon is not responsible for manufacturing all these products, the company and its web site do provide unprecedented access to them in terms of availability and lower cost, with little substantial information regarding the products' origin or value, and no information regarding the multi-dimensional, short, medium, and long term impacts of its purchase.

Information visualizations (both static and interactive) use some type of graphical representation to display items in a collection. For example, *The U.S. Gun Deaths* project (Kirk, Kois, and @GunDeaths) uses curved, moving, coloured lines to display length-of-life projections for people who were killed by a gun (see Figure 1). The *Out of Sight, Out of Mind* project (Pitch Interactive) also uses curved, moving lines, but this time to display drone strikes on Pakistan (see Figure 2). In both designs, data took an abstract yet metaphorical form, though with very different rationales: (1) line length as time passing and line length as distance travelled; (2) color and shape as life lived vs. life that could have been lived (orange and white) and strike vs. life lost; and (3) position on screen as life arch vs. burial.

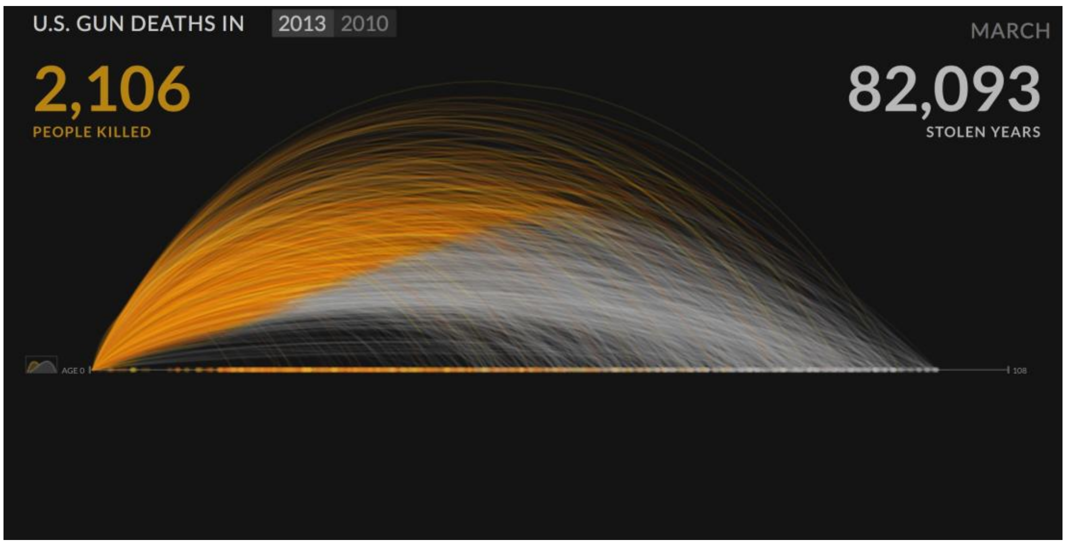

Figure 1: U.S. Gun Deaths (Kirk, Kois, and @GunDeaths).

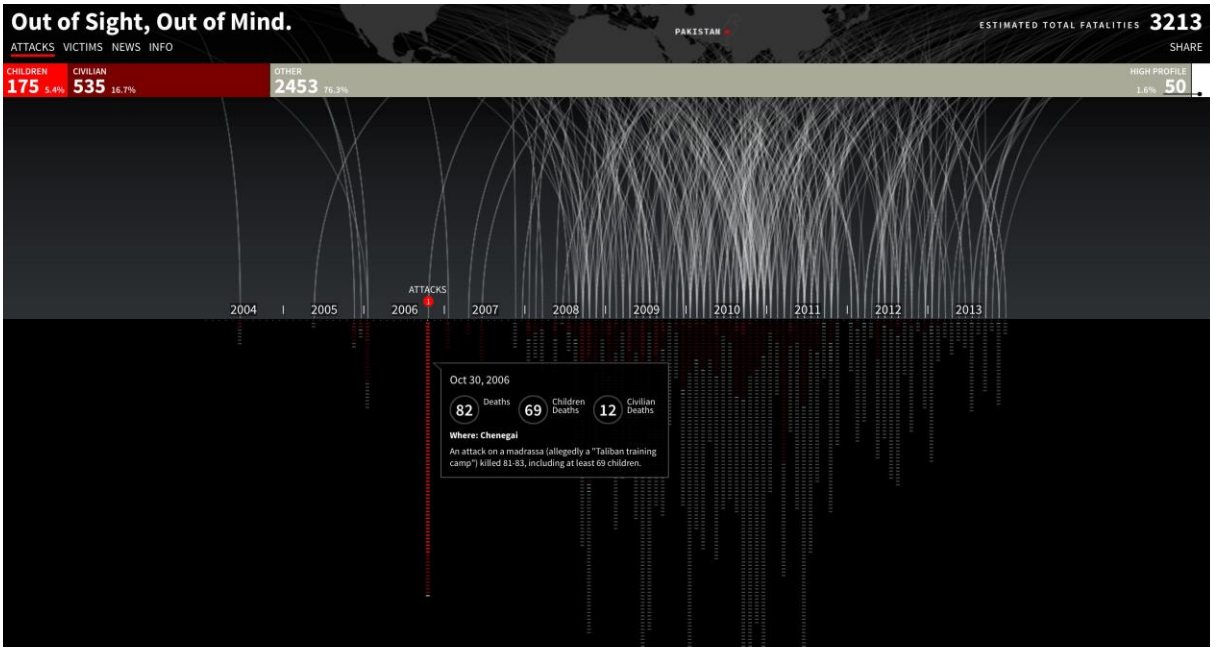

Figure 2: Out of Sight, Out of Mind (Pitch Interactive).

Building on past work in rich-prospect browsing, feminist HCI, and text visualization, this paper is a theoretical reflection on the political nature of graphically-represented data, and aims to explore three questions: (1) What constitutes meaningful representation of items in a collection;
(2) What are the potential implications (consequences) of different kinds of representations; and
(3) What is the value of sameness or difference in graphical representation?

Relevant Literature

In data visualization, designers make a choice, first, what type of graphical object – in Peirce’s terms an icon, index, or symbol – will stand in for the items found in a particular collection. Second, they consider the details of form. Such representation may, in some instances, be closely evocative of the collection item’s original form (e.g. a thumbnail photo may stands in for a larger version of the photo), or the form may be arbitrarily assigned (a line stands in for a person’s life span). The *Mandala Browser,* for example, (see Figure 3) shows a macro view on the entirety of *Romeo and Juliet,* with each speech in the play displayed as one, color-coded dot (http://mandala.humviz.org.) When a user clicks on a dot, she can read the corresponding speech in a text frame located to the right of the browser. The graphical representation (dot) of each speech is arbitrary. Using Peirce’s terms, this interface uses symbols as graphical representations of data: dots and Shakespeare have no natural connection to one another.

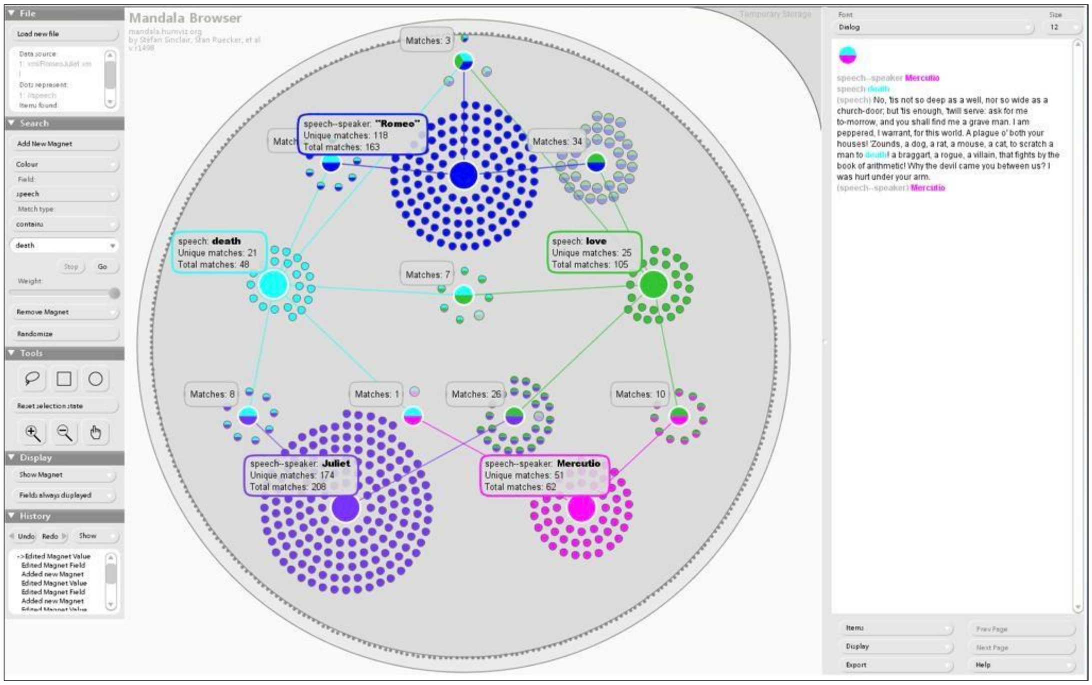

Figure 3: The Mandala Browser. Design by Constanza Pacher and Sandra Gabriele.

In another example, the *Paper Drill* interface displays an article's citations (items in that collection) using a gridded series of coloured squares, with each square representing one citation source and the colour the centrality of the article in terms of what it cites and how often it is cited by the other papers (Ruecker et. al., 2011). The squares have been chosen by the designer to represent a "heat map" of articles in the collection. In this case the connection between a square, its colour, and the text is indexical—defined by some sensory feature. In the case of the *Paper Drill*, the colour of the square is representative of the level of "hotness", or popularity, of that paper within the collection. Popularity is determined by how many citations the papers it cites get and how many papers also cite it.

The additional benefit of the representation found in the *Paper Drill* interface, is that the small geometric shapes allow for a prospect view on all the items in a very large collection (Figure 4 shows 1,666 citations by 36 authors).

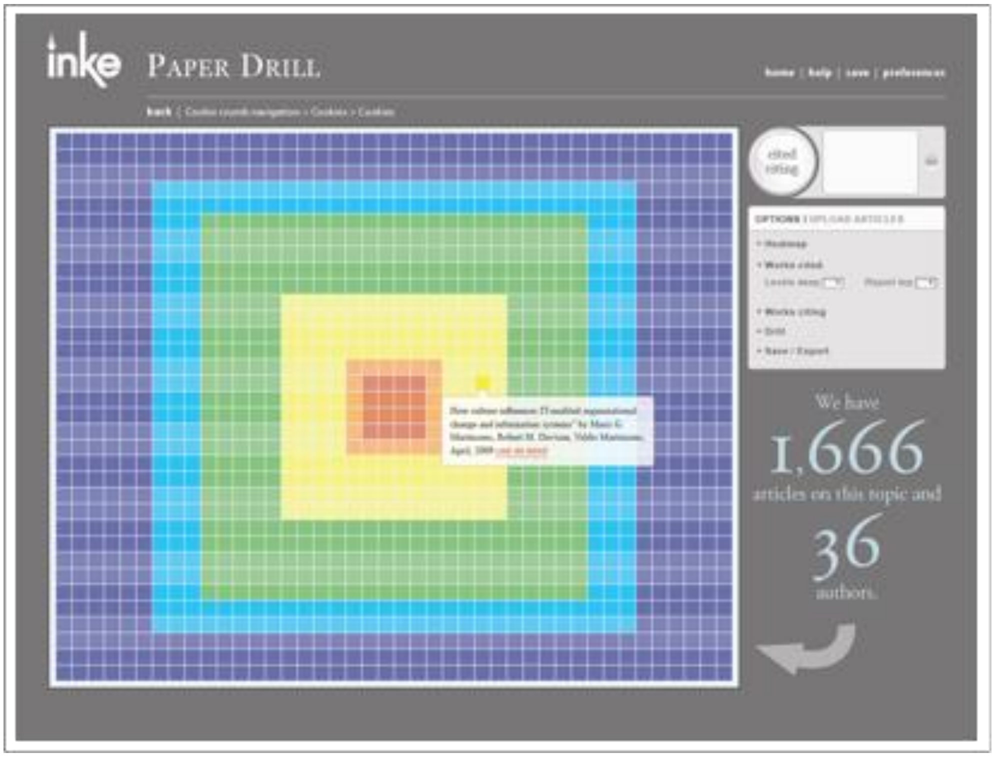

Figure 4: The *Paper Drill* by Milena Radzikowska.

For some collections, on the other hand, an iconic item representation may be not only possible, but preferable given the nature (and size) of the collection and the types of tasks that are likely to be performed with it.

In the *Slot Machine* interface (see Figure 5) the first column features the entirety of Gertrude Stein's *Making of Americans* (Stein, 1995) in micro text, with subsequent columns generated based on a user's search of a repeated phrase. All columns are aligned along a reading slot that magnifies the repeated phrase and its immediate context. The *Slot Machine* is a good example of an interface where there is a close connection between the item in the collection (the novel) and its graphical representation (a column made of a micro text version of the novel) (Radzikowska, et. al., 2007). In this case, the close connection in representation (Peirce's icon) is of benefit since it facilitates an exploration of the text across and within multiple contexts. At the same time, a micro text representation of a novel (even one that is not 1,000 pages in its original form) requires the use of specialized technology, such as a wall-size display, making such representation possible but impractical (Radzikowska, et. al., 2007).

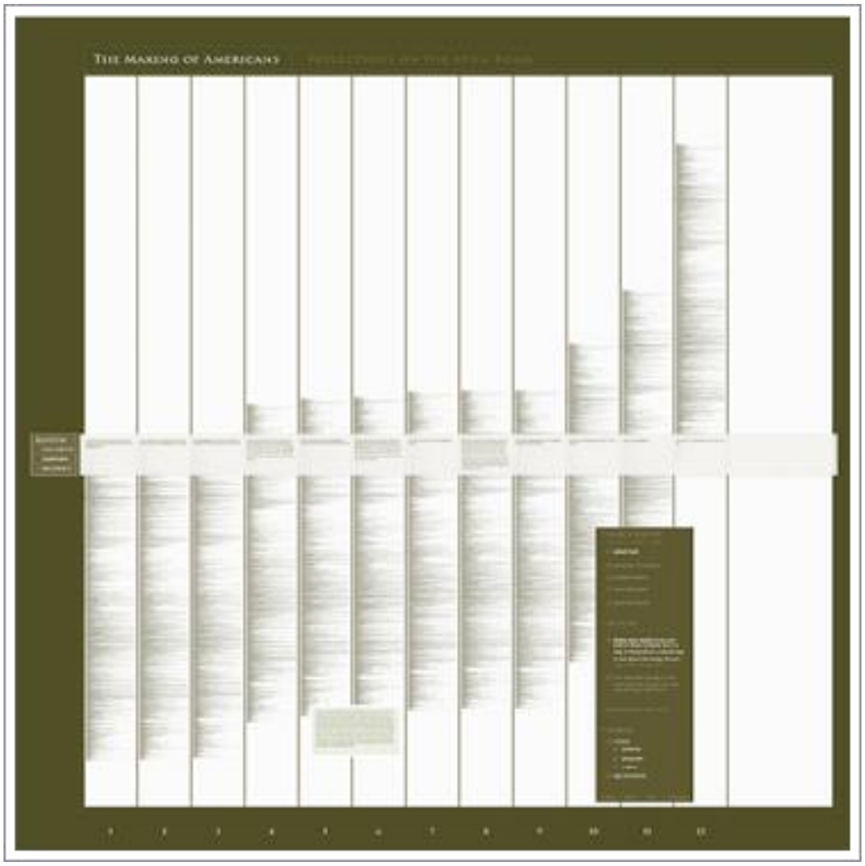

Figure 5: The Slot Machine by Milena Radzikowska.

In our previous work on rich-prospect browsing (Ruecker et al., 2011), we began discussing ways that the user's current task may benefit from a dynamic change in the representation of the items. For example, being able to rotate an object in order to see it from all sides (Peirce's icon) might be crucial for some tasks, while a simple text label (Peirce's index) for the same object might be sufficient for others, and for still other uses a dot (Peirce's symbol) might be enough. Our argument is that the interface should do everything possible to provide the user with options for manipulating a display showing all items in the collection, since it is not possible for the designer, or the user, to know in advance all of the tasks that the user will want to perform.

Research Methods

To interrogate graphical data representation we must first accept three premises: (1) that a designed artefact can hold some kind of argument; (2) that design thinking and making contribute to human knowledge; and (3) that designed artefacts are complex agents that act upon our world.

In design the creation of an artefact can be, in and of itself, a way to formulate an argument about designing similar artefacts. This idea, proposed by Galey and Ruecker (2010), is based on what they see as theoretical affinities between design and book history scholarship, experimental interface design sharing much in common with the emerging practice of peer-reviewing digital objects in scholarly contexts. Both design and book history engage in interpretation and in making and both, if properly contextualized, “can contribute to a theoretical framework for new questions”(Galey and Ruecker, 2010). Similarly, Bardzell argues that a design can function as a form of research. He builds on work in aesthetic cognitivism, and asserts that design, in its ability to tell us something about reality, contributes to human knowledge (“Design as Inquiry”). The critical analysis of designed artefacts “positions us in a potent space between the past and the future. Failing to recognize design as a hermeneutic process means failing to understand how our inherited cultural record actually works” (Galey and Ruecker, 2010). If we are to conduct a thorough critical analysis, we must consider designed artefacts at the macro, meta, and micro levels—each artefact is (1) a collection of multiple, designed parts; (2) a totality that is something far more complex than the additive nature of its individual components; and (3) is inherently contextual and context-dependent (Radzikowska, 2015). That design matters, that it exists, that it will have an impact, and that this impact can be for the positive has, more recently, entered the multi-disciplinary discourse. Fuller (2008) acknowledges that “objects, devices, and other material entities have a politic—that they engage in the arrangement and composition of energies, allow, encourage or block certain kinds of actions” and writes that “these concerns have also more recently been scrutinized by the interdisciplinary area of science and technology studies”(7).

Discussion

1. Meaningful Representation

What does it mean for graphical representation to be meaningful? Perhaps surprisingly, Peirce’s categories are orthogonal to this question. Document icons, for example, though graphical are often not very meaningful, especially in the context of other, similar icons—when used to represent hundreds or even dozens of documents, they become “a complex pattern composed of identical elements” (Ruecker, 2003). They convey the size of the collection—I have several dozen annotated articles on feminism—but very little information about the unique characteristics of the items contained within the collection (see Figure 6).

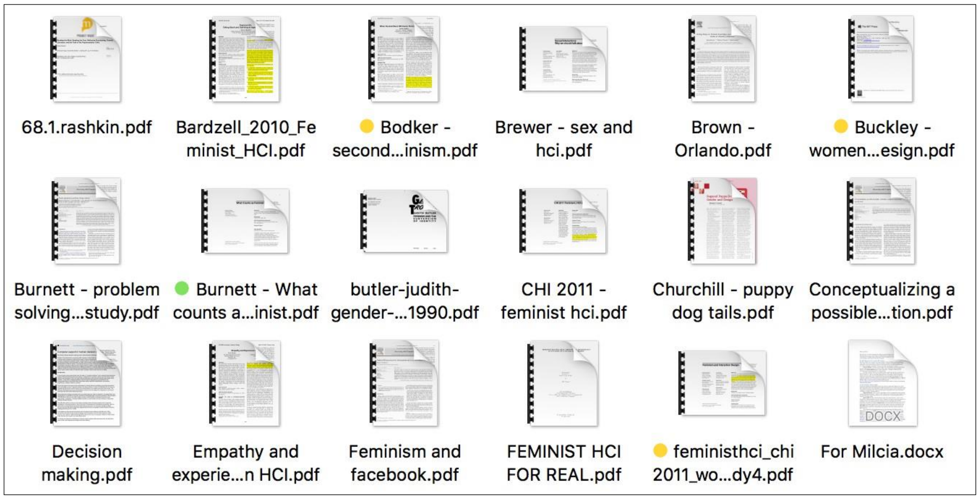

Figure 6: In a file folder, textual labels need to be added to differentiate one file from another; suggesting that the graphical representation alone is insufficient. In fact, filenames are a required component.

What constitutes meaningful graphical representation depends, primarily, on what kinds of knowledge we expect to extract from its browsing. For example, one shopping cart icon is meaningful by itself, signifying online purchasing. In a collection of shopping cart icons, we may be able to explore the different ways icons have been drawn to represent online shopping, their colour structures, that some are carts and some bags (see Figure 7).

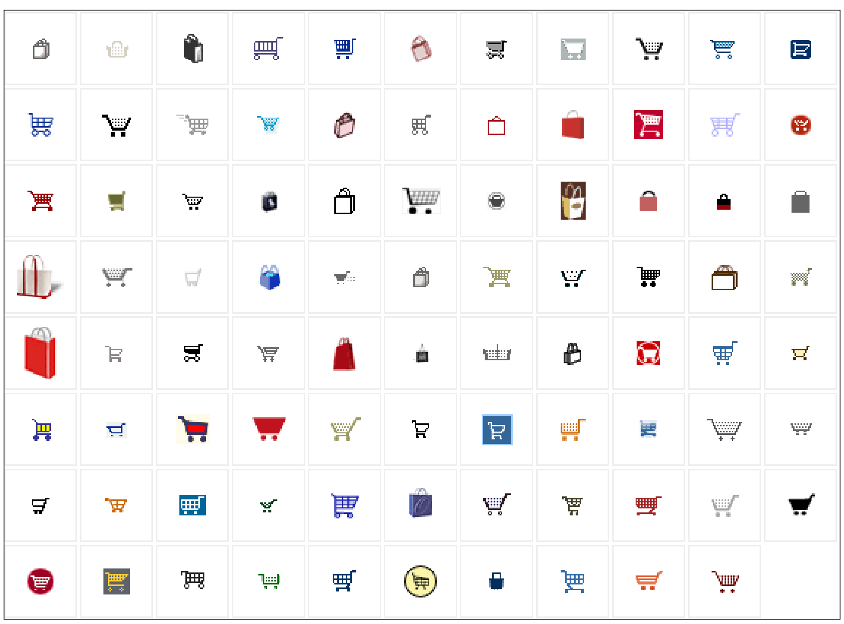

Figure 7: Collection of shopping cart and bag icons

(http://www.intersmash.com/300images/images/carts.gif.)

If we have additional information attached to each graphical representation—country of origin, type of online shop where it is used, attributes of its designer—our exploration can become much more meaningful. We may be able to consider cultural or social trends, change in design over time, and much more. In another example, imagine that we are looking at images of sheep, from a farm that produces merino wool to be made into sweaters. If our hope is to see photos of Betsy (one specific sheep) and her human and animal family but we are, instead, shown generic sheep- like icons, we will be disappointed. If, instead, we are exploring sheep types found in New Zealand's wool industry, illustrations that are based on but that abstract sheep categories may be more informative.

The design of meaningful graphical objects to represent collection items (in data visualizations, for example) must balance the following: amount of useful information that can be communicated through the object's graphical form, meaningful graphical difference between individual or groups of items, and restraint in form complexity to allow for the simultaneous display of numerous collection items at a small size. How we interpret difference and sameness and, more importantly, whether we attach hierarchical value to the emergent categories, may determine whether we place emphasis on one set of data over another.

Whether we consider a graphical representation meaningful, may or may not have anything to do with its visual complexity. For example, an abstract line may have intended meaning (life lived) and emotional impact (life prematurely cut short); while a photograph (which is far more complex in terms of visual information) may simply intend to say faculty member (see Figure 8). We must, though, differentiate between intended and unintended meaning. In the *Faces of Innovation* example as shown in Figure 8, the intent of the designer may have been to simply display MRU's researchers and make them recognizable to others through the use of portraits instead of abstract head icons. However, by looking at the entire collection of portraits, we may be able to conclude that MRU has a predominantly Caucasian population. Depending on the context of viewing (as a potential hire, for example) such information may be quite meaningful.

Figure 8: *Faces of Innovation* displays all researchers at MRU (Mount Royal University).

2. Implications & Consequences

In our recent work on visualization for decision support in the Alberta oil industry, the results provided to the user (the data that becomes part of a collection and, subsequently, a visualization) are highly dependent on the specific choices of variables, which in this context are referred to as constraints[1]. Traditionally, the choice of constraints that can be manipulated by the user consists of processes, parts, and materials. For example, one constraint might consist of the number of trucks, while another is the number of barrels of refined oil. Figure 9 shows one design solution to the problem of graphical representation: a gear shape represents connection (a cause and effect relationship between parts); a solid gear represents a constraint; and a pie chart gear represents the solution to a decision problem. Thus, in Figure 9 we can see four constraints at play to generate a solution, and a separate solution (possibly made in the past) that isn't currently under consideration (the non-geared pie chart). Coloured gears signify liquid ingredients, while coloured and textured gears signify dry ingredients. This distinction provides an additional layer of information and a greater range of options for colour coding.

[1] Our industry partner asked us to use ice cream manufacturing instead of oil refinement and extraction as our scenario. This is reflected in the models and design prototypes (as shown in Figure 8) generated for the project.

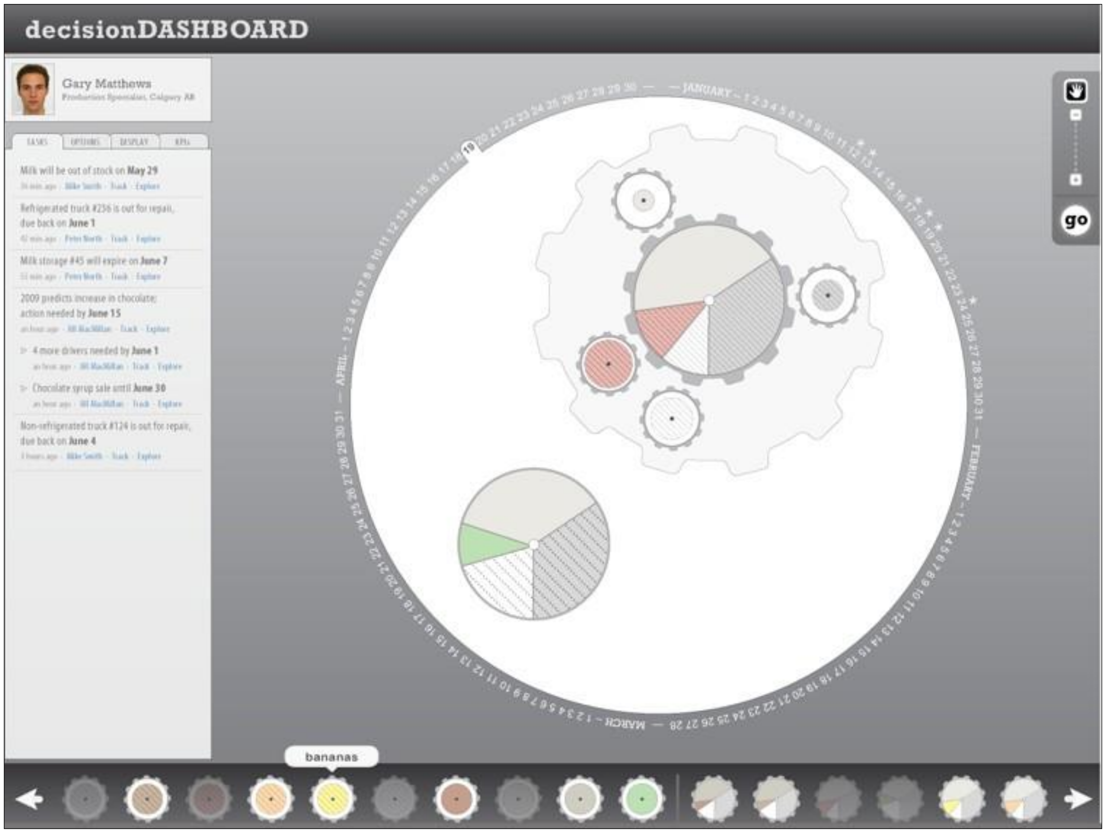

Figure 9: Gears represent ice cream recipes and smooth circles represent ingredients (Radzikowska et al., 2011).

The gear shapes are reminiscent of the inner working of clocks, car parts, and machines, representing factory production or industry. Further yet, Symbolically, the gears represent connectivity and interdependence, since a gear's teeth make it possible for it to connect with other gears in order to form a more complex system. Thus, the fact that in this design the manipulation of one gear affects or changes another constitutes a rather short metaphorical leap—a reasonable mapping to the natural world. Furthermore, in the specific instance of the case study presented by our industry partner—decision making in manufacturing—gears are part of a familiar, if rather dated, visual language. This design leverages what has been recently termed skeuomorphism: where an artefact retains ornamental design cues from structures that were necessary in the original (Basalla, 1989). Adding the soft colour palette to the gears—pinks, yellows, blues, and tans—is meant to suggest a particular type of machine-based production: ice cream flavours. This particular colour choice moves the design away from the harsh contrasts typically associated with metal gears and industrial machinery: greys, blacks, and dark browns. The pastel colour palette has a potential downside, anchoring the design in ice cream production, making it less transferable into other types of manufacturing; however, if the communication goal was to make oil extraction and refinement appear less threatening (softer, friendlier, more people-friendly), then the colour palette may be spot on.

3. Value of Sameness vs. Difference

Manufacturing decisions (particularly, but not exclusively, those within the oil industry) appear to have consequences—both positive and negative—on individuals, communities, and environments. Currently, the oil industry does not consider constraints that have to do with human and environmental factors as part of its decision making process. If it did, we may have to consider graphical representation for non mechanical constraints: individuals, groups, communities, working spaces, and natural and constructed environments and ask ourselves:

- should human / environmental constraints should be given the same graphical form as mechanical constraints, such as raw materials, waste, resulting products, and methods of distribution;
- do constraints such as working hours and conditions, sick leave, and safety concerns require a different graphical treatment; and
- should people constraints be treated as visually different from environmental constraints?

If all these constraint types are given the same graphical representation, would the design suggest that people are considered of the same importance as the amount of waste or its disposal? Sameness and difference have many potential interpretations. If all items in a collection are treated the same way graphically, despite any difference in terms of data they may have, they are likely to be seen as equal. Equality may be seen as a positive in contexts where there has been, traditionally, a devaluing of certain categories over others. However, it is also possible that, in the manufacturing sector for example, equality of mechanical, human, and environmental constraints may be seen as a devaluation of the human condition.

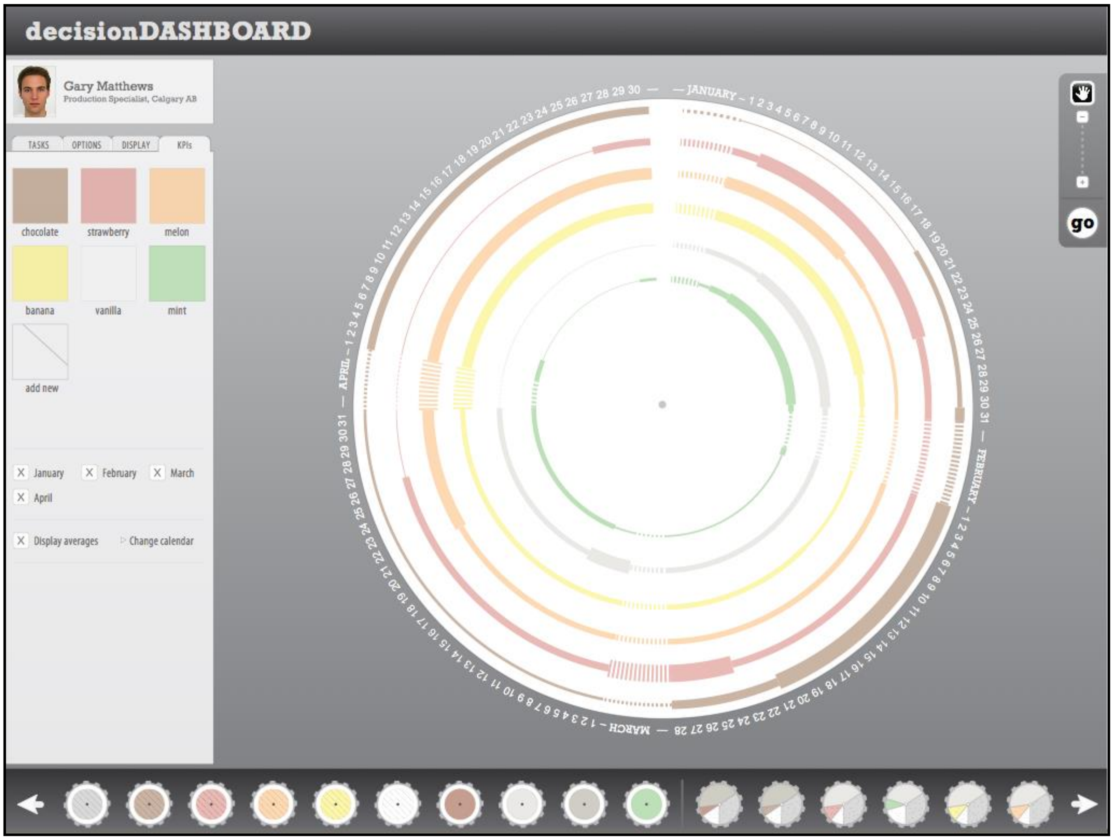

Figure 10: An example of arbitrary representation: production values are represented by varying thicknesses of line, and constrained by a circular calendar (Radzikowska et al., "Human Decisions for a Machine World", n.pag.).

In certain cases of graphical representation, each item in a collection needs to have a different appearance in order to differentiate it from the others, but each one is meant to be seen as equal in significance. In figure 10, for example, each ingredient is given its own line and colour. They area means to be viewed as the same in terms of importance, just unique in terms of type.

If each category of items within a collection is given a perceptually different graphical representation, users would be able to see that the data set consists of, in some way, unique groupings. In certain data visualization, difference may hold more information than sameness, when that difference is, in fact, meaningful. And in certain other data visualizations, difference may, in fact, be critical to its functionality. Consider the *Pill Browser*, as an example (see Figure 11): it displays the actual photographs of 1000 prescription pills available on the US market. The purpose of this browser is to identify rogue pills by their appearance—imagine you found a loose pill on your Grandpa's floor and you're trying to figure out which of his many medications it is. The *Pill Browser* lets you conduct a visual search based on the shape and colour of the pill. Each pill has to be different in appearance (and in its graphical representation within the collection) for it to be useful in terms of accurate identification[2].

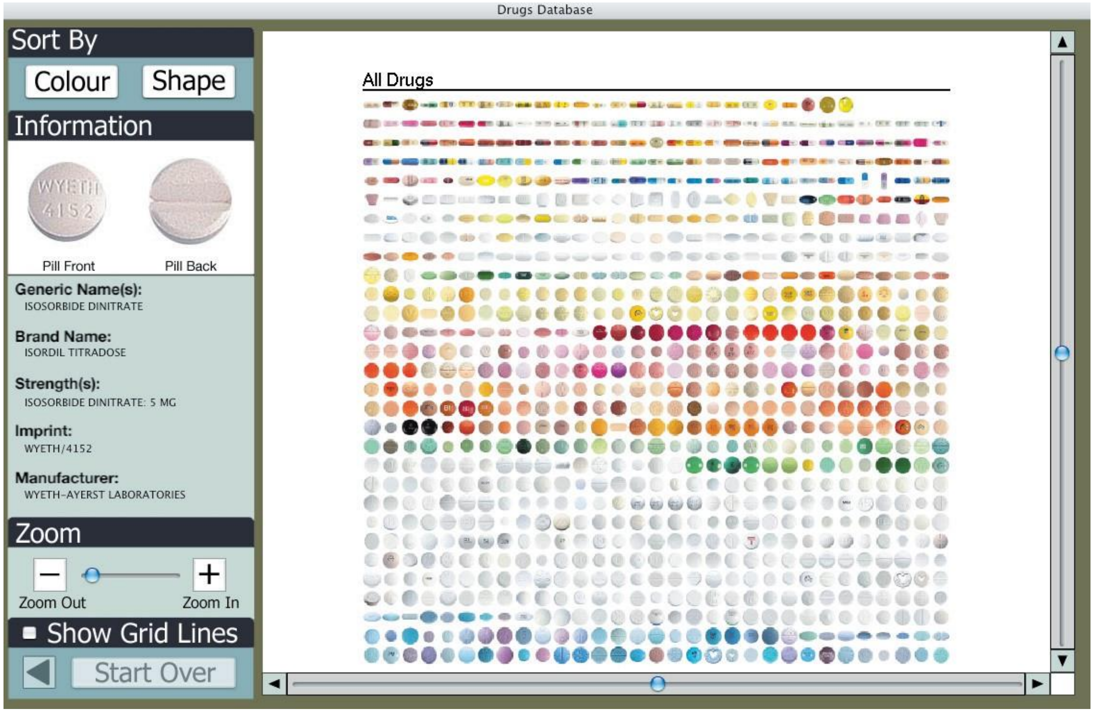

Figure 11: The Pill Browser (Ruecker, Given, Sadler, and Ruskin).

Difference may indicate a hierarchy of importance. If our choice of graphical representation places emphasis on one category of constraints over another, that emphasis will also hold meaning. For example, in the NYTimes visualization *What do you think is the most important problem facing the country today?* colour differentiates category types, and difference in size per category comparatively quantifies concern (the larger the rectangle, the higher the level of concern) (see Figure 12).

[2] Many pills are, in fact, visually different from each other, and the designers of the Pill Browser leverage that difference in this data display. They also show that many other pills are simply round and white, suggesting that accurate identification may be difficult if not impossible.

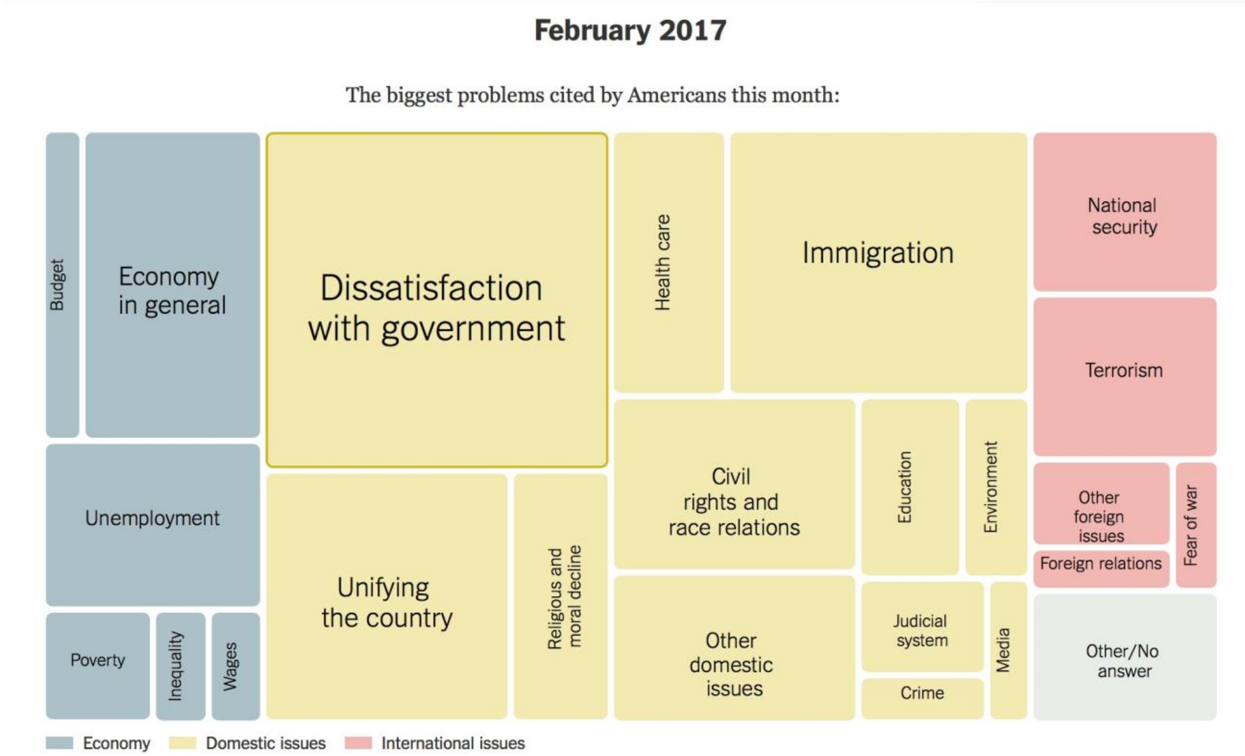

Figure 12: *What do you think is the most important problem facing the country today?* from 2016, the Trump presidency (https://www.nytimes.com/interactive/2017/02/27/us/politics/most-important-problem-gallup-polling- question.html?_r=0).

Conclusions & Future Work

The kinds of graphical representation selected for a data visualization may impact the user's perception of a category (and the data set it contains), the category's level of significance and, potentially, the user's, and in some cases the category member's, perception of the self. The first part of the preceding statement makes sense to most designers: that the form we assign to an object will impact how that object is interpreted. In data visualization, however, the design challenge can be far more complex: interpretation must often be balanced with functionality. While we can display a thousand dots on the screen, each representing a different item within a collection, our current, common denominator display technology cannot do the same for photographs with enough size and detail for the content of each photograph to become identifiable as a unique item with unique characteristics. This is particularly true if the entire collection is meant to be viewed and browsed as a complete set. This is true in the *U.S. Gun Deaths* and the *Out of Sight, Out of Mind* projects, and the *Mandala* and *Pill* browsers.

Another option is to place the form of representation under the control of the user, so that someone interested in photos could have them represented as thumbnail images, while another user, in the context of a different task, could have them represented by typological categories such as type of camera, time of day, location, subject matter, and so on. Previous studies (e.g.

Giacometti et al., 2009) have suggested that instead of permanently changing representation of the individual items, a preferred strategy is to temporarily change it, then change it back, and use the recently visible information instead as a means of organizing the items. If a user is assigning a graphical representation to her own identity (as is demonstrated through the personalization power of FaceBook, Twitter, and Instagram profiles), she may be empowered by that act; however, if she is assigning a representation to another person or community, such an act may be disempowering or marginalizing.

Finally, since our focus in this paper is on representations of data rather than functions, we are not concerned with buttons, scrollbars, and other navigational objects. However, the distinction between data and functionality is often blurred, both in the case of desktop icons where double-clicking a document will open up the program that is associated with it and, in the case of many data visualizations, where the graphical representation of data becomes an access point for more data. In effect we need to extend our understanding of graphical representation beyond the visual images themselves, into the data, code, or data + code that are connected to the images.

References

Basalla, George. *The Evolution of Technology*. Cambridge, UK: Cambridge UP, 1989.
Print. Chandler, Daniel. *Semiotics: The Basics*. Routledge; 2 edition, 2007. Print.
Fuller, Martin. *Software Studies: A Lexicon*. Cambridge, MA: MIT Press, 2008. Print.
Galey, Allan, and Stan Ruecker. "How a Prototype Argues." *Literary and Linguistic Computing* 25.4 (2010): 405–424. Print.
Giacometti, Alejandro. The Texttiles Browser: An Experiment in Rich-Prospect Browsing for Text Collections. Thesis, University of Alberta, 2009. Print.
Papanek, Victor. Design for the Real World: Human Ecology and Social Change. NY: Bantam, 1973. Print.
Radzikowska, Milena. *The Paper Drill*. Graphic. Stan Ruecker and the INKE Research Group. “The Paper Drill.” Digital Humanities 2011. 19–21 June 2011, Stanford, CA. Presentation. n.pag.
---, Stan Ruecker, Walter Bischof, Michelle Annett, and Fraser Frobes. “Gearing Up: Visualizing Decision Support for Manufacturing.” *Journal of the 2009 Chicago Colloquium on Digital Humanities and Computer Science* 1.1 (2009): n.pag. Web. 5 July 2015.
---, Stan Ruecker, Carlos Fiorentino, and Piotr Michura. “The Novel as Slot Machine.” Annual Conference of the Society for Digital Humanities (SDH/SEMI). 28–30 May 2007, Saskatoon, CA. Presentation.
Ruecker, Stan. Affordances of Prospect for Academic Users of Interpretively-Tagged Text Collections. Diss. University of Alberta, 2003. Print.
---, Lisa Given, Elizabeth Sadler, and Andrea Ruskin. “Building Accessible Web Interfaces for Seniors: Similarity Clustering of Pill Images.” Include 2005. 5–8 Apr. 2005, London, UK. Presentation.
---, and the INKE Research Group. “The Paper Drill.” Digital Humanities 2011. 19–21 June 2011, Stanford, CA. Presentation.
---, Milena Radzikowska, and Stefan Sinclair. Visual Interface Design for Digital

Cultural Heritage: A Guide to Rich-Prospect Browsing. London, UK: Ashgate, 2011. Print.
Stein, Gertrude. The Making of Americans: Being a History of a Family's Progress. 1925. Normal, IL: Dalkey Archive Press, 1995. Print.

Author Biography

Milena Radzikowska, MDes, PhD

Dr. Milena Radzikowska is an Associate Professor with current research interests in the areas of decision support, feminist HCI, humanities visualization, critical design, and information design. She designs, teaches, and conducts research as a feminist, a committed mentor, and community builder. Her work in human-computer interaction is reciprocally informed by her passion for creating safer, more inclusive and compelling spaces, both digital and analog. Her design and research work is interdisciplinary and collaborative, marked by a passion to work in the service of others.

Over the past 15 years Dr. Radzikowska worked on over a dozen projects designing human-computer interfaces with researchers from the digital humanities, primarily exploring interface design for decision support; visualization for large text collections; online support environments for breast cancer survivors; and wildlife tracking systems for provincial parks. Her work has been iterative and experimental – meant to challenge existing interface design conventions and explore unique alternatives to complex visualization problems. It comes out of the belief that design can and should aim to make a difference to individuals, communities, and society at large. She has co-authored more than 50 publications on data visualization, aesthetics, interaction design, interaction theory, and design for large text collections. She is also the co-author of the book *Visual Interface Design for Digital Cultural Heritage: A Guide to Rich-Prospect Browsing.*

Stan Ruecker, MDes, PhD

Dr. Stan Ruecker is the Anthony J. Petullo Professor in Design at the University of Illinois. His research is radically interdisciplinary and international in scope. He has been tenured at the University of Alberta, the globally renowned Institute of Design in Chicago, and most recently at the University of Illinois.

Professor Ruecker has co-authored with over 230 different people, working collaboratively in over 20 academic disciplines. His work is broadly in the design of information and communication. More specifically, his research explores the implications of new media on communication and interpretation, including speaking, writing, and experience.

Dr. Ruecker's research has focused for the past 20 years on the future of reading, producing more than two dozen experimental prototypes. He is currently exploring physical interfaces for complex conceptual work, such as text analysis, modeling time, and designing experience. He is also part of an international group developing new predictive models of key concepts for use by designers. Their current topic is how design can help encourage people to expand from forming opinions to holding multiple interpretations.

As the Anthony J. Petullo Professor in Design, Dr. Ruecker’s focus is on raising the profile of design research at the University of Illinois and expanding the role of design research globally.