

Exploring Growing Information Spaces

Marian Dörk, Sheelagh Carpendale, Carey Williamson

Department of Computer Science, University of Calgary
Calgary, Alberta, Canada

{mdoerk, sheelagh, carey}@ucalgary.ca

ABSTRACT

Growing information spaces bring not only informational challenges, but also new opportunities for exploration. Drawing from maps as an analogy, we conceptualize information seekers as explorers of expanding information landscapes. While text has been the dominant frame of reference for information-seeking interfaces, visualization has the potential to expose faceted overviews and provide many ways of seeing information. Considering recent technological trends on the Web, many new forms of information seeking are feasible that utilize rich semantics and interactive graphics. With these developments on our side, we advocate a shift from information needs and relevance to more open-ended exploration and discovery of growing information spaces.

1. INTRODUCTION

Today's information spaces are growing at a tremendous pace, which can be attributed to both organizational and individual increases in data generation, collection, and sharing. Thousands of status updates, blog posts, photos, videos, and many other types of resources continually appear on the Web. While this growing abundance of information can be viewed through the lens of information overload, we see growing information spaces as opportunities for exploration and discovery. However, existing navigation techniques provide limited support for getting a sense of what is 'out there'.

With the growth of the Web, the techniques for traversing it have also evolved. At the Web's beginning, one followed hyperlinks from one page to the next to find interesting, but not always relevant, information. With the arrival and evolution of search engines, people gained more control over the information they access by entering keywords and going through result lists. Now, people also monitor information sources by subscribing to their favourite websites and following short status updates of friends and acquaintances. These forms of information access are useful in many ways, yet, they imply a certain narrowing of information pipelines centred around our personal search terms, subscriptions, and

social network. It is less about learning what is 'out there', and more about looking into a mirror constructed by keyword relevance, expressed interests, and social popularity.

We anticipate that the next major shift in information navigation will be one from textual to visual, from relevance to overview. This shift is foreshadowed by considerable research on visualization for search over the last 15 years [e.g., 2, 3, 11, 13] and the wide-spread use of tag clouds and interactive maps on the Web. Conventional search interfaces still lack visual overviews of results let alone entire information spaces. Only small sets of items are displayed and ranked by search term relevance. While the value of overviews to represent information collections is recognized [9], search interfaces employ the visual modality primarily for previews of individual resources using thumbnails or text snippets. To provide information seekers with meaningful overviews of and orientation within growing information spaces, we propose a visual paradigm that borrows concepts from cartography and advocates a multiplicity of perspectives.

Considering their general role for humanity, maps can serve as a valuable analogy for making sense of information spaces. There has been quite a bit of work on visualizing non-spatial information using metaphors from geography, for example, document landscapes [16] and domain maps [1]. However, it has been argued that cartographical metaphors taken too literal can cause misunderstandings and confusion [8]. For the purpose of this paper, we are less interested in maps as concrete sources of inspiration for visual design. Rather on a conceptual level, we seek to explore the role of maps for making sense of the world and examine commonalities for exploring growing information spaces.

Maps represent geospatial regions and landmarks at different scales. The represented phenomena can be natural (such as mountain ranges) or man-made (such as cities). While topographic maps are created to provide accurate representations of physical characteristics, other map types, such as thematic maps, can represent any information that has a spatial aspect such as political boundaries or economic indicators. When we use a map it becomes easier for us to make sense of our environment and confidently engage with it—physically or intellectually. The great utility of maps is then twofold, as they help us to think about the world by supporting a sense of overview, and to act within it by supporting a sense of orientation. It is important to note that it is the human who is our primary concern. Therefore, we want to emphasize *thinking about and acting within the world* as dual purposes of maps, and explore the parallels for interactive visualizations of information spaces.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

HCIR 2011, Mountain View, USA.

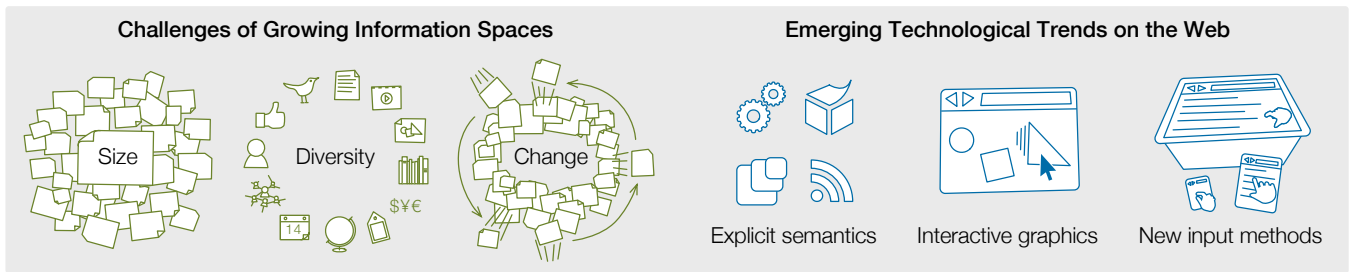


Figure 1: Information challenges and technology trends drive and enable new forms of information seeking.

2. CHALLENGES AND OPPORTUNITIES

Technologies such as compasses, sextants, chronometers, and celestial navigation allowed seafarers find their way across oceans to encounter new lands. While this comparison may sound corny to some, we are approaching a similar age of exploration. This time, the challenge is to discover expanding information landscapes and harness the opportunities of emerging technological trends on the Web (see Figure 1).

2.1 Growing Information Spaces

Like new worlds to be discovered, growing information spaces have unique characteristics related to their size, diversity, and dynamics that challenge us to find new ways of making sense of them.

While the Web’s size is hard to measure, consider over a million book reviews on the community site LibraryThing, more than 500 million profiles of active members on Facebook, and more than 5 billion photos on Flickr. Simple statistics are not sufficient to make sense of these collections.

The wide spectrum of Web resources translates into a large variety of facets that could be used to explore information. Common facets across many resource types are tags, date/time of publication, and location. One could also think of other more specific facets, such as price and rating for commercial products. The challenge is to extract relevant information, design expressive visualizations, and build scalable systems that expose this multitude of facets and present them based on the interests of information seekers.

The explosion of content shared by people on the Web also constitutes a tremendous information change. For example, more than 10 million article edits are executed each month on Wikipedia, more than 100 million status updates are submitted to Twitter every day, and almost 50 hours of video is uploaded to YouTube every minute. These levels of dynamics hold interesting, yet largely untold stories about current trends and historical developments.

2.2 Technological Trends on the Web

Similar to the advances in maritime navigation enabling the ‘discovery’ of foreign lands, we see emerging trends on the Web as enabling technologies for a more visual and exploratory form of information seeking. These trends centre around ongoing and projected developments with regard to semantics, interactive graphics, and input methods.

There are more and more data on the Web that are well structured and semantically organized. These semantically enriched resources are increasingly exposed in search engines as corresponding result snippets. Photos and status updates are associated with location information providing geospatial context for these resources. The ongoing rise of seman-

tically structured information on the Web poses a great opportunity for new visual exploration interfaces.

At first, information visualization on the Web relied on browser extensions, but now increasingly rich graphics are becoming natively supported by browsers as Web standards are maturing. The upcoming HTML5 specification includes browser-native APIs for embedding rich multimedia contents and interactive graphics without any use of third-party plugins. Bitmap, vector, and 3D graphics will likely be supported by all the main browser providers. While it will take several more years until these developments reach the majority of Web users, innovative research and design can now prepare with these trends in sight.

Considering that the Web is accessed more and more via cellphones, tablets, and laptops, the form factors and input methods are changing as well. The increasing use of touch-enabled smartphones such as the Apple iPhone and Android phones allows for novel ways of interacting with the Web. Beyond touch, these mobile devices also have more implicit input methods such as device location and orientation that can be exploited for new types of interaction. Large touch-enabled wall displays and interactive tabletops bring new opportunities for interacting with information in new settings such as the public sphere and private homes. Collaborative and casual forms of information seeking can be supported with these ongoing and projected advances in mind.

3. A NEW FRAME OF REFERENCE

To find new ways to think about and interact with growing information spaces, it can be helpful to examine where the map analogy breaks. The main difference between cartography and visualization is that a map generally has geographic space as the primary frame of reference: positions on a map correspond to locations in the world. There is no comparable frame of reference for digital information spaces. Instead, we face a multitude of resource types and facets that is continually expanding and evolving.

3.1 Textual Flattening

While the sheer volume and variety of information on the Web may defy any attempt to be reduced into one frame of reference, in practice, search engines have established text as the dominant mode of representation. Even faceted search interfaces represent rich information aspects as lists of text-based links or controls [10, 17]. However, similarly to how the doctrine of the Flat Earth constrained the perception of the world, the current primacy of text influences how we view the information spaces around us. Resources are indexed based on textual contents or metadata, then queried using textual search terms, and finally presented mostly as

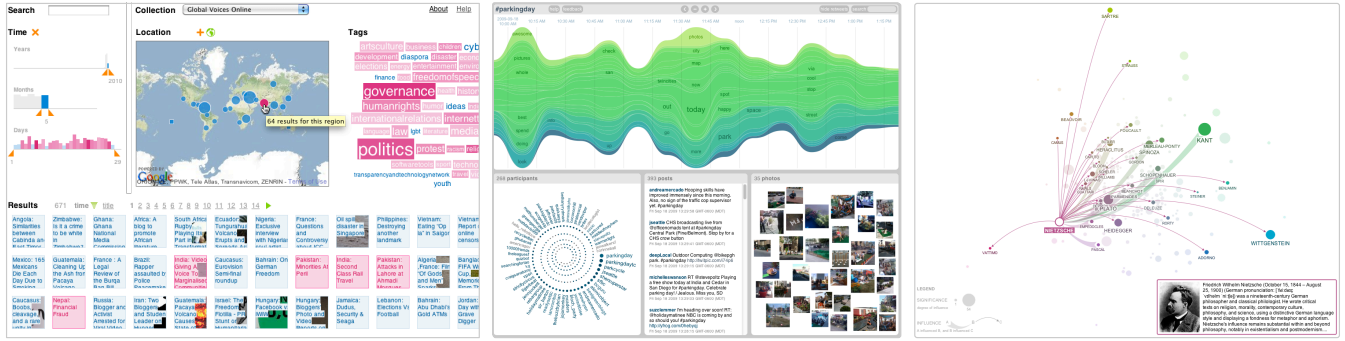


Figure 2: Visualizations of blog posts (left), tweets during events (middle), and people (right) are designed to help information seekers explore growing information spaces and discover patterns and relationships.

text-based, ranked search results. This approach makes it easy to find something specific, however, it also transforms multifarious information spaces into flat worlds of text.

Text is computationally convenient as keyword indices, and ranked result lists are much easier to generate than visual information displays. However, text does not make the best use of our visual perception and it limits the ways we engage with information. In particular, search interfaces rely on the separation of information access between query formulation and result navigation as evidenced by the ubiquitous search box and result list. In a sense, this query-response paradigm inherits the ancient legacy of the all-knowing oracle that one would ask the right question and then interpret the response. In a recent commentary in *Nature*, Etzioni argues for the perfection of this trend by transforming search engines into general question-answering machines [7]. However, more exploratory forms of information seeking are not motivated by explicit questions or problems, but rather by general interests and entertainment [14, 15]. While current search interfaces provide tiny glimpses into information spaces along individual search terms, we never see their entirety. This has implications on how we think about and interact with information spaces.

3.2 Information Landscapes

Building on emerging technological developments, interactive visualizations are powerful instruments that provide us with new ways of seeing and exploring information. Like compasses on a map allow the seafarer to envision a voyage at a high level, visualization may enable us to traverse information with a sense of overview and orientation. Coupled with more sophisticated data mining and richer semantics, interactive visualizations can reveal otherwise hidden facets and expose them as new frames of reference.

In contrast to text, which is used today as a unified perspective on information, visualization offers room for many alternative ways of seeing information. Information visualizations can convey meaningful overviews that show distributions of resources across multiple facets and allow the exploration of relationships and resources. The goal for visualization design then is to expose important facets not only as a way of looking at information spaces, but also a tool that invites interacting with them. Using examples from our own research, we explore how visualizations can be seen as maps of growing information spaces, in that they enable both reflection and exploration.

For example, VisGets as shown in Figure 2, left, offer multiple coordinated views, that provide visual overviews along time, location, and tags situating resources in the context of their collection. In addition to visualizing blog posts along these facets, VisGets allow the viewer to interactively formulate multidimensional queries by dragging temporal sliders, zooming and panning to a region on a map, and selecting tags of interest [3]. Many other facets are feasible by integrating query formulation and visualization.

Different from the stable nature of the physical world, growing information spaces are by definition in flux. For example, during large events, the many status updates posted on Twitter constitute a highly dynamic information space. To explore how topics in such information streams evolve over time, Visual Backchannel provides a topic visualization that represents the ebb and flow of topics within thousands of tweets (see Figure 2, middle). The interface also provides self-adjusting views of participants and photos, and allows the viewer to apply filters along several facets [6].

Using visualizations, we can both survey how resources are scattered across a range of facets and then use interaction techniques as steering wheels to direct our exploration along our interest and curiosity. Multiple facets can also be integrated to expose new relationships that were not visible before. Information about people’s inspirations and interests is a good case in point, which illustrates that data often contains both explicit and implicit relations. For example, in order to represent significant personalities such as philosophers one can consider both their influences and similarities. By combining multidimensional scaling with graph drawing, we can encode similarity as position and influences as curved edges (see Figure 2, right). The resulting visual formations of EdgeMaps represent the extent of influence across interests and allow the viewer to explore philosophers along similarities and influence connections [4].

3.3 Exploration and Discoveries

In the past, the boundaries of people’s imagination about the world were pushed by daring explorers and skillful cartographers (among others). To help people expand their perspective on growing information spaces, we need to consider not only tools and technologies, but also people’s traits and tendencies. We propose a positive persona inspired by the flaneur wandering the arcades of 19th century Paris. Instead of framing information seekers around needs, problems, and deficiencies, we emphasize their inclinations towards curios-

ity, creativity, and reflection, and suggest explorability as a new design principle for information-seeking interfaces [5].

In contrast to the emphasis on relevance and accuracy forming the basis of conventional search interfaces, information seekers often have vague and changing information interests. Similar to the sea explorers not knowing what they would find, information seekers often do not know what to look for and what they may encounter. Rather than locating specific resources, information seeking along these lines satisfies higher-level information needs such as learning and discovery [12]. Visualizations can help people orient themselves in growing information spaces and discover new and possibly unexpected resources. Towards this end, the combination of direct manipulation and visual representation provides powerful ways of situating data items in the context of their larger collection. While traditional visualization aims to help domain experts make findings and gain insight based on thorough data analysis, there is great potential for visualization to help casual information seekers to make discoveries of interesting resources and relationships.

4. CONCLUSION

The trend towards more refined information access through question-answering, subscriptions, and social networks may help us to find factual answers and filtered resources, but does not promise a better sense of overview and orientation of what is out there. It's almost like giving us the coordinates of a city without the corresponding map. In contrast, flipping through the pages of an atlas allows us to imagine travelling the world and visiting foreign countries. Likewise interactive visualizations can make large information spaces accessible both intellectually and practically. Considering the emergence of semantics and graphics on the Web and the surge of touch-enabled devices, interactive visualizations have the potential to turn abstract information spaces into graspable information landscapes. We think that exploring information spaces using interactive visualizations can be a more holistic and exploratory form of information seeking.

5. REFERENCES

- [1] Katy Börner, Chaomei Chen, and Kevin W. Boyack. Visualizing knowledge domains. *Annual Review of Information Science and Technology*, 37(1):179–255, January 2005.
- [2] Edward C. Clarkson, Krishna Desai, and James D. Foley. ResultMaps: Visualization for search interfaces. *TVCG: Transactions on Visualization and Computer Graphics*, 15(6):1057–1064, Nov/Dec 2009.
- [3] Marian Dörk, Sheelagh Carpendale, Christopher Collins, and Carey Williamson. VisGets: Coordinated visualizations for web-based information exploration and discovery. *TVCG: Transactions on Visualization and Computer Graphics*, 14(6):1205–1212, Nov/Dec 2008.
- [4] Marian Dörk, Sheelagh Carpendale, and Carey Williamson. EdgeMaps: Visualizing explicit and implicit relations. In *VDA2011: Conference on Visualization and Data Analysis*, page 78680G. IS&T/SPIE, 2011.
- [5] Marian Dörk, Sheelagh Carpendale, and Carey Williamson. The information flaneur: A fresh look at information seeking. In *CHI '11: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1215–1224. ACM, 2011.
- [6] Marian Dörk, Daniel Gruen, Carey Williamson, and Sheelagh Carpendale. A visual backchannel for large-scale events. *TVCG: Transactions on Visualization and Computer Graphics*, 16(6):1129–1138, Nov/Dec 2010.
- [7] Oren Etzioni. Search needs a shake-up. *Nature*, 476(7358):25–26, 2011.
- [8] Sara I. Fabrikant, Daniel R. Montello, and David M. Mark. The natural landscape metaphor in information visualization: The role of commonsense geomorphology. *Journal of the American Society for Information Science and Technology*, 61(2):253–270, 2010.
- [9] Stephan Greene, Gary Marchionini, Catherine Plaisant, and Ben Shneiderman. Previews and overviews in digital libraries: designing surrogates to support visual information seeking. *Journal of the American Society for Information Science*, 51(4):380–393, 2000.
- [10] M.A. Hearst. UIs for faceted navigation recent advances and remaining open problems. In *HCIR 08: Workshop on Human-Computer Interaction and Information Retrieval*, 2008.
- [11] Marti A. Hearst. TileBars: Visualization of term distribution information in full text information access. In *CHI '95: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 59–66. ACM, 1995.
- [12] Gary Marchionini. Exploratory search: From finding to understanding. *Communications of the ACM*, 49(4):41–46, 2006.
- [13] Anselm Spoerri. InfoCrystal: A visual tool for information retrieval. In *VIS '93: Conference on Visualization*, pages 150–157. IEEE, 1993.
- [14] Ryen W. White, Bill Kules, Steven M. Drucker, and m.c. schraefel. Supporting exploratory search: Introduction. *Communications of the ACM*, 49(4):36–39, 2006.
- [15] Max L. Wilson and David Elsweiler. Casual-leisure searching: the exploratory search scenarios that break our current models. In *HCIR 10: Workshop on Human-Computer Interaction and Information Retrieval*, 2010.
- [16] James A. Wise, James J. Thomas, Kelly Pennock, David Lantrip, Mark Pottier, Anne Schur, and Vern Crow. Visualizing the non-visual: Spatial analysis and interaction with information from text documents. In *InfoVis 1995: Symposium on Information Visualization*, pages 51–58. IEEE, 1995.
- [17] Ka-Ping Yee, Kirsten Swearingen, Kevin Li, and Marti Hearst. Faceted metadata for image search and browsing. In *CHI '03: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 401–408. ACM, 2003.