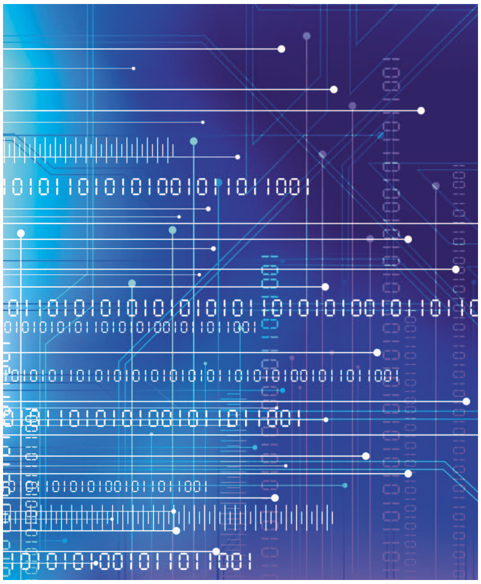

Ubiquitous Analytics: Interacting with Big Data Anywhere, Anytime

Niklas Elmqvist, *Purdue University*

Pourang Irani, *University of Manitoba*

Ubilytics amplifies human cognition by embedding analytical processes into the physical environment to make sense of big data anywhere, anytime.

With more than 4 billion handheld devices in use in the world today, mobile computing is quickly becoming the universal computational platform (P. Baudisch and C. Holz, "My New PC Is a Mobile Phone," *XRDS: Crossroads*, vol. 16, no. 4, 2010, pp. 36-41). Building on this new wave of portable devices are inherently mobile personal computing activities such as microblogging, social networking, and photo sharing.

Mobility is now propagating to more professional activities such as data analytics, which are no longer restricted to the workplace. In fact, the rise of big data increasingly demands that users have access to data resources anytime, anywhere to support decisions and activities such as travel, telecommuting, and distributed teamwork.

In other words, in the realm of data analytics, Mark Weiser's vision of ubiquitous computing is on the threshold of becoming a reality ("The Computer for the Twenty-First Century," *Scientific Am.*, vol. 265, no. 3, 1991, pp. 94-104). In particular, the quiet revolution in mobile computing has led to the embedding of massive infrastructures of computational power in our everyday surroundings that are just waiting to be harnessed.

UBILYTICS

We refer to the use of multiple networked devices in our local environment to enable deep and dynamic analysis of massive, heterogeneous, and multiscale data anytime, anywhere as *ubiquitous analytics*, or *ubilytics*. This approach not only liberates analysts from the confines of the office and leverages computer hardware now mostly used for personal and entertainment purposes, but also taps into deep cognitive processes that could improve the analytical process.

As Figure 1 shows, *socially distributed cognition* models human thought as an intrinsically system-level process that goes beyond sensory-input-based information processing in the brain to involve interactions with the body, the surrounding physical world, tools and artifacts, and other humans (E. Hutchins, *Cognition in the Wild*, MIT Press, 1995).

Current personal computers provide only small viewing portholes and limited input bandwidth into the digital universe, whereas embedding interactive analytics components into smartphones, tablets, large displays, interactive tabletops, and even head-mounted displays (such as Google Glass) enables a more natural, distributed approach to interacting with big and complex data.

MOTIVATING EXAMPLE

Consider a public health scenario in which authorities are trying to contain a pandemic outbreak in a local area.

In this real-time emergency situation, data comes in from multiple

 Published by the IEEE Computer Society

heterogeneous sources, including historical information from existing databases as well as dynamic feeds from hospitals, doctors' offices, pharmacies, and so on. Furthermore, the data has many forms, such as admissions data, health records, sales statistics, and so on. Unexpectedly useful insights might also come from social networks ("Who interacts with whom?") and credit card data ("Who has gone where?").

All of this information is temporal and spatial in nature, and all of it must be considered to determine where the outbreak started and how it's spreading. Finally, agents, analysts, and decision makers in this rapidly evolving scenario are distributed in both time and space.

Figure 2 shows a conceptual view of a ubilytics approach to supporting this scenario in which an ensemble of networked devices—ranging from smartphones and tablets for field agents, to laptops and desktops for individual analysts, to large displays and tabletops for collaboration—constitute a single reasoning environment.

Key stakeholders such as scientists, doctors, health officials, epidemiologists, and policymakers can collaboratively explore the data and, drawing on their collective expertise, make conclusions, verify them, and take appropriate action. The diversity of the participants' skills and backgrounds, and the use of interactive analytics embedded in the physical environment, will lead to decisions that are superior to those made by any individual.

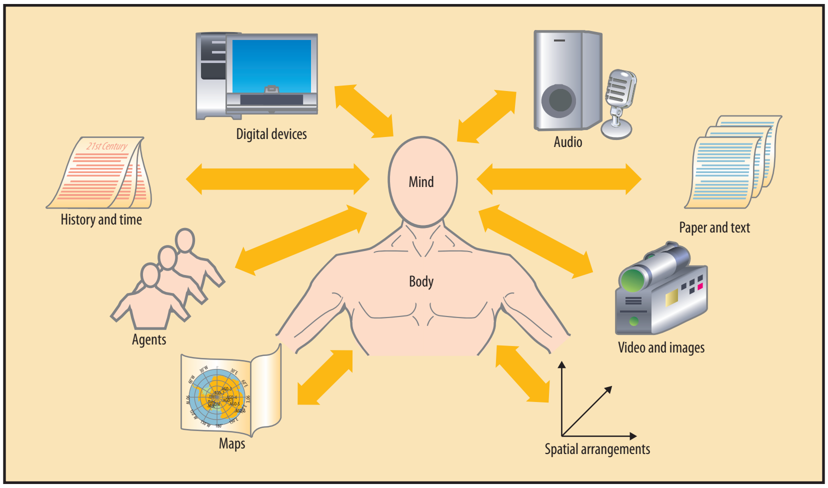

Figure 1. Socially distributed cognition. Human cognition can be seen as a system-level process involving not only the brain and its sensors but also physical space, tools and artifacts, and other people.

Figure 2. Ubilytics approach to managing a pandemic outbreak. In this scenario, multiple stakeholders use a menagerie of devices to collaborate using a shared data repository.

CHALLENGES

To make ubilytics possible, researchers must address several challenges.

Networked device ensembles

Traditional data analytics applications are built for a single computer. While such software might use concurrency to leverage multiple cores, distributed computational platforms are required to enable seamless analytics across an ensemble of networked devices, each of which could join or leave the shared environment at any time.

Limited computational resources

While mobile computing capabilities are steadily increasing, handheld devices still have orders of magnitude less computing power and memory compared to standard computers. This significantly impacts mobile applications, particularly analytics applications that carry out statistical, cluster, or probabilistic analyses and often incorporate big data, complex algorithms, and rich and interactive visual representations.

Polymorphic hardware

Supporting effective analyses in a ubiquitous setting requires fully exploiting the hardware capabilities of each participating device and accommodating its weaknesses. Mobile computing in particular is plagued by the tradeoff between minimizing device form factor and

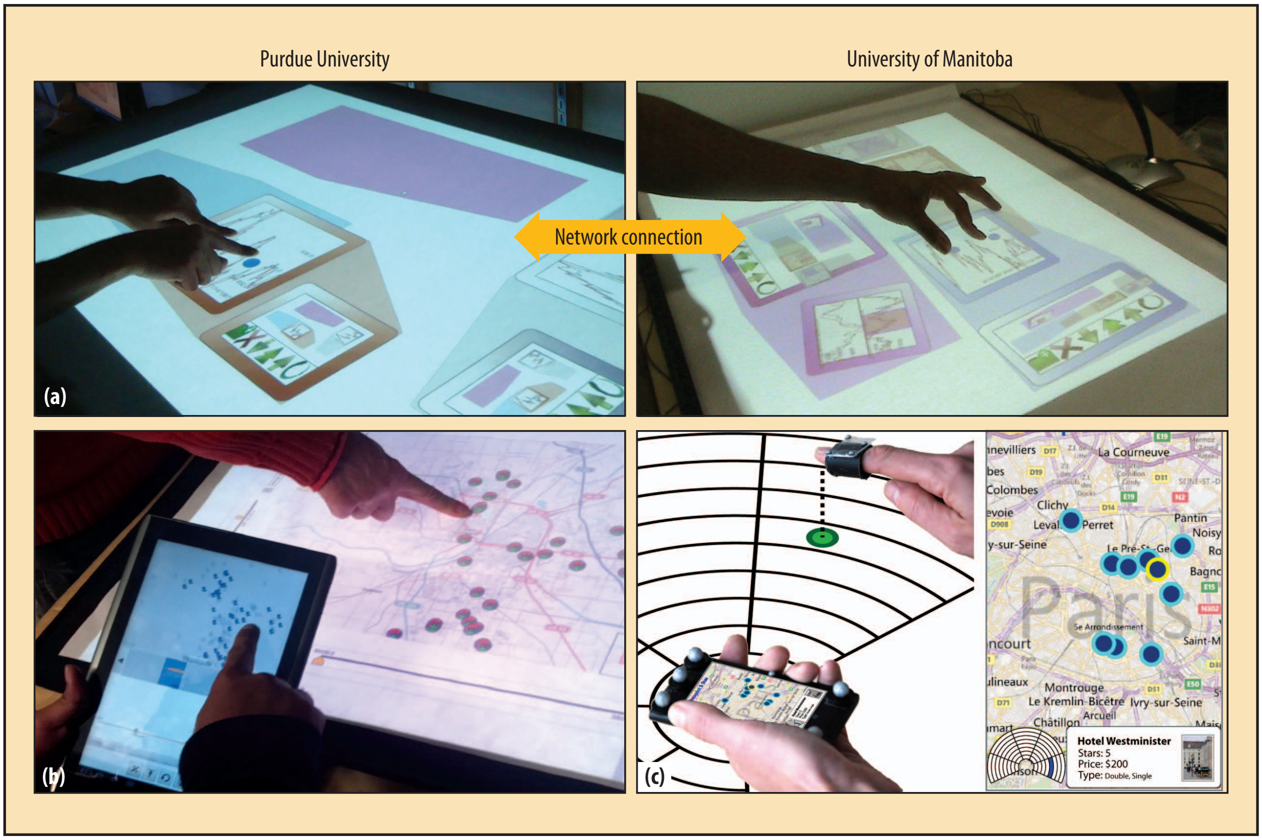

Figure 3. **Three example ubilytics systems. (a) Collaborative data analysis using the Hugin mixed-presence toolkit. (b) Tablet and tabletop for collaborative search using Branch-Explore-Merge. (c) Appropriating physical space around a mobile device using Around-Device (AD) binning.**

maximizing interaction and display surface. Even though current smartphones consist almost entirely of a multitouch display, these screens are still very small and require good eyesight and precise input, making them poorly suited for spatially intensive tasks.

Connecting groups of users

Individual analysts will remain the core user group for most analytics software, but efficiently tackling many big data problems requires correspondingly large teams. Groups contribute additional viewpoints, broader expertise, and multiple roles and authorities to the analytical process, which often leads to better results than those achieved by individual analysts. However, connecting collaborators across space and time, promoting awareness and consensus among team members, and resolving conflicts can be problematic.

UBILYTICS EXAMPLES

We have attempted to address all of these challenges in our own work designing, building, and evaluating three separate ubilytics systems: the Hugin toolkit, Branch-Explore-Merge, and Around-Device (AD) binning.

Hugin toolkit

The Hugin toolkit is a mixed-presence framework for supporting collaboration scenarios in which some of the participants are colocated in the same physical space and others are connected over a network (K.T. Kim et al., "Hugin: A Framework for Awareness and Coordination in Mixed-Presence Collaborative Information Visualization," *Proc. Int'l Conf. Interactive Tabletops and Surfaces* [ITS 10], ACM, 2010, pp. 231-240). Figure 3a shows the Hugin connecting two multitouch tabletops in our research groups' respective labs, allowing geographically separated analysts to visually explore various multidimensional datasets together.

Hugin uses a scalable client-server architecture to connect multiple heterogeneous tabletops using a platform-independent network protocol. To accommodate different display sizes, the virtual space for each tabletop is stacked and scaled to fit the maximum physical size. Finally, the toolkit has native support for group awareness and conflict resolution, including mechanisms for visualizing remote touches, overview maps, and access control.

Branch-Explore-Merge

Conflicts and interference can arise even when analysts collaborate in the same physical space and at the same time. The Branch-Explore-Merge protocol applies ideas from source code revision control systems such as CVS and SVN to the analysis process by allowing individual analysts to branch off a current query, explore new parameters, and then merge new findings into the shared state (W. McGrath et al., "Branch-Explore-Merge: Facilitating Real-Time Revision Control in Co-Located Collaborative Visual Exploration," *Proc. Int'l Conf. Interactive Tabletops and Surfaces* [ITS 12], ACM, 2012, pp. 235-244). Figure 3b shows our BEM prototype being used to search for real estate in the Greater Lafayette, Indiana, area.

The prototype combines Android smartphones and tablets with a digital tabletop. The Android devices serve as clients that connect to the BEM server running on the tabletop. Costly computation and rich visualization are offloaded to the server since it runs on a personal computer, whereas the clients have limited computational resources. Such a collaborative platform facilitates fluid transitions between shared and personal workspaces.

Around-Device binning

Traditional mobile interfaces aren't designed to handle big or spatially intensive data. Thus, as mobile devices become the primary means for interacting with large datasets on the go, novel interfaces are needed to fully support the sensemaking process.

Mobile interfaces can be evolved to support big data interaction tasks using minimal mobile sensing augmentation. AD binning is our initial concept for enabling users to interact with large search content without requiring minute and tedious operations, such as multiflicking on a limited-size display, as is common today (K. Hasan, D. Ahlström, and P. Irani, "AD-Binning: Leveraging Around Device Space for Storing, Browsing and Retrieving Mobile Device Content," *Proc. 31st Int'l Conf. Human Factors in Computing Systems* [CHI 13], ACM, to appear, 2013).

As Figure 3c shows, a mobile device can use AD binning to sense interactions beyond its screen. Users can interact with objects, either by placing them into spatial bins and later retrieving them, or by having the system automatically place items in midair and around the device.

Such augmentations create new opportunities for transparent interactions with large data repositories. More importantly, they mitigate the spatial limitations that mobile devices impose. This general approach has also been dubbed *appropriated interaction* because it borrows space from the surrounding physical environment to offset small display and input sizes (C. Harrison, "Appropriated Interaction Surfaces," *Computer*, June 2010, pp. 86-89).

A ubiquitous approach to analytics introduces a paradigm shift in sensemaking methodology. Analysts will be able to interact with big and complex data on an individual basis or collaboratively, on the go or in the office, synchronously or not.

Our own research only provides a glimpse of the full potential of ubilytics. Much more work is necessary to lay the groundwork for this growing and increasingly important discipline.

Acknowledgments

The work of Niklas Elmqvist was partially supported by US National Science Foundation award IIS-1253863. The work of Pourang Irani was partially supported by the Canadian NSERC Discovery program. Any opinions, findings, and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the respective funding agencies.

Niklas Elmqvist *is an assistant professor of electrical and computer engineering at Purdue University. Contact him at elm@purdue.edu.*

Pourang Irani *is an associate professor of computer science at University of Manitoba, Canada. Contact him at irani@cs.umanitoba.ca.*

Editor: Naren Ramakrishnan, Dept. of Computer Science, Virginia Tech, Blacksburg, VA; naren@cs.vt.edu