Beyond the Desktop Metaphor in Seven Dimensions

Thomas P. Moran and Shumin Zhai
IBM Almaden Research Center

To be published as a chapter in the book
Designing Integrated Digital Work Environments: Beyond the Desktop Metaphor
Victor Kaptelinin and Mary Czerwinski (Editors)

The ubiquitous use of the desktop metaphor as the primary means of interacting with information is perhaps the earliest, and arguably the most profound, landmark of user interface design. Ironically, such a success is both a great past achievement and a difficult future challenge to overcome. Computing technologies and user experiences available to people in our current web-driven world are evolving rapidly. In fact, the strict concept of the desktop metaphor is already a "straw man" notion, but it can help us characterize where we were and where we are going. We are already in mid-flight from the desktop metaphor to somewhere else. Although we cannot be sure where we are going, we can discern different dimensions in which things are changing.

The research presented in the chapters of this book represents some notable efforts in moving beyond desktop-metaphor-based computing. In this concluding chapter we reflect and comment on seven dimensions of change along which we see future integrated digital work environments being different, as experienced by users, from today's computing environment. Our analyses and speculations are based on the chapters in this book and our own research, as well as the HCI literature and information technology trends in general. In the spirit of concluding this book, we do this in very broad strokes that try to capture major themes.

Here, in a nutshell, are the dimensions of change that we will examine:

1. The basic change is that personal information is being liberated from the constraints of the desktop/office metaphor. It is being dispersed in the networked world in what we might call a "personal information cloud".
2. Several other kinds of changes follow from this. The desktop metaphor standardized, and thus limited, the ways information was presented. New ways of organizing personal information are spawning a great variety of new representations and visualizations.

3. The desktop metaphor was designed for a standardized computational form factor, the workstation and laptop. The proliferation of new forms of computing devices both requires and exploits the information cloud to allow information to "follow the user".
4. The desktop metaphor is built around keyboarding and pointing. The multiplicity of devices of different sizes and functions forces designers to develop new modes and modalities of physical interaction techniques.
5. Not only is information liberated from the desktop, but so also are software applications. Functional computations delivered as services from servers make these functions available independent of specific devices.
6. The desktop metaphor creates a personal office isolated from others except through limited channels. More and more personal information clouds are intersecting in richer ways to make collaborating with others participating in large-scale social communities easier.
7. The desktop/office metaphor creates an arena focused on a variety of generic office tools geared to low-level interaction tasks. Future computational work environments should be centered around the meaningful activities that people are engaged in, which requires an explicit representation of the concept of activity in the information cloud.

Note that these seven dimensions are not exhaustive; there are dimensions of change, such as moving from rigid to more adaptive representations; but these seven seem most related to the body of work exhibited in this book. In what follows we reflect and comment on each of these dimensions, relate them to each other and to the chapters in this book, and conclude with a brief review on where we stand on these dimensions.

Dimension 1: From the office container to the "personal information cloud"

The desktop metaphor was originally invented to support office work. The metaphor is really a personal *office* metaphor. The metaphorical desktop itself is a display screen with various office-relevant objects – documents (overlapping windows), folders (icons), and tools (e.g. printer icons) – in a freeform arrangement. There is also a metaphorical file system organized as a hierarchy of folders and files, sort of like file cabinets. Further, there is a metaphorical mail-based inbox, providing ways for messages and attached documents to enter and leave the office.

The dominant feature of the desktop/office metaphor is that information is *contained* in the office, in both a cognitive and a physical sense. Users understand that information objects have a place: on the desktop, in a folder, in the inbox, etc. But there is also a physical reality to the containment notion – the digital information is actually stored in the physical memory of the personal computer. The metaphor enables the user to understand and manage the information in the computer's physical store.

There is an ongoing trend to interact with information outside the metaphorical office. Workers in business settings have for a long time been using file servers to retrieve, backup, share, and archive information. The world-wide web has made remote information accessible within the metaphorical office. Much information does not have to be stored in the office machine for it to be readily available. And people are not just retrieving, but also putting information on the web. Millions of people use hosted email services. The evolution of the web to "Web 2.0" is enabling people to not only retrieve, but also to create personal content and annotations on the web. So, personal information such as email is now commonly stored outside of the office machine.

But there is also a deeper cognitive trend in the way users understand how to manage their information. There is great cognitive comfort in the idea of *containment* – that a document is contained in some folder, a known *place* where it is located and can be found. The desktop/office metaphor is based on the notions of containment and place. But these notions are being eroded by the ability to effectively search for information, first on the web and now on the desktop/office itself. Users do not have to be concerned about where information is if they can effectively get at it by search.

We do not believe that search will be the only method to get at information. There are strong individual differences in relying on search. For example, some people do not create email folders at all, and rely on searching their inbox. But other people are "frequent filers" (Whittaker & Sidner, 1996). There are good reasons to pay the cost of manually structuring information, such as organizing and planning benefits (Jones, Phuwanartnurak, Gill, & Bruce, 2005).

Structuring information does not require containment; it only requires *reference* – the ability to create descriptions that can reference information objects. While this is inherently a more abstract notion than containment, people are gaining experience with the concept every day in using the web. The web emphasizes references (links) between pages and de-emphasizes the notion that the information is contained in places (but it does not totally eliminate the notion of places, i.e., servers).

There is an interesting analogy between information and money. Money can also be kept in a place (at home or a bank), or it can be placeless. Although there was also a great deal cognitive comfort in keeping money "under the mattress," eventually most people gave up such a comfort and accepted the fact their money is dispersed inside of financial institutions, which in turn loan and invest the money all over the world. It is almost unknowable where each individual's money precisely is. All that matters is that they can get it or transfer it on demand.

As users disperse and "destructure" their personal information, there is less need for the desktop/office metaphor to be the organizer of the information. We believe that the metaphor is being replaced by more abstract and sophisticated organizers, based on over a decade of experience by millions of people with information technology. Thus let us use the term *Personal Information Cloud* to refer to the "working set" of information that is relevant to the individual and his work. We are not promoting this term as a profound new notion; it is just a convenient label for our use here. We do not think that a "cloud" is a particularly useful metaphor for either users or designers. In contrast to the desktop metaphor, which was consciously designed by the first user interface designers, the personal information cloud will probably not be "designed" at all, but rather will evolve as a set of organizing principles based on the collective experiences of developing and using the web.

This general personal information cloud is what people need to interact with, not with a particular device or metaphor; the latter are mediators of this interaction. There are several requirements to the personal information cloud to be useful:

1. *Personal*. It should contain most if not all information that is relevant to the individual and his activities.
2. *Persistent*. It should be preserved.
3. *Pervasive*. It should be always accessible from a variety of devices, programs, and services, i.e., it "follows the individual".
4. *Secure*. The information should be secure and private at an appropriate level. This is a significant issue when information is not held locally (although having information locally is not in itself assurance of privacy in a networked world).
5. *Referenceable*. Each information object in the cloud should ideally have a unique ID (or permalink) and support a protocol for retrieval.
6. *Standardized*. The information needs to be in standard formats to that it is usable by a variety of devices, programs, and services.

7. *Semantic*. The cloud should be based on an extensible scheme of semantically-rich metadata, so that it can be understood by a variety of programs and services in different contexts.

Many of the other dimensions follow naturally from this notion of a personal information cloud: new information representations, new device form factors, and new interaction techniques. Social interactions and activity management can also be better enabled by a personal information cloud.

Dimension 2: From the desktop to a diverse set of visual representations

The most noticeable feature of today's personal computing environment is its visual interface, which is based on the desktop metaphor, and a set of GUI (graphical user interface) rules and conventions to represent information objects and regulate interaction behavior. As a virtual world the "physics" of the conventional desktop to some extent resembles the real world, including constant scale, continuity, fixed place (of file location), and "Newton's first law" ("an object at rest stays at rest until acted upon by force" or "objects on the desktop stay where the user places it.). Today's desktop computing environments also organize information hierarchically into files in folders. Most computer users have lived digitally in this virtual world for more than a decade.

Much of this book is devoted to issues such as how successful is today's desktop interface, how users really use it (Ravasio and Tscherter, this volume), and, especially, what alternative representatives there are (Freeman's chronological Lifestreams representation, Karger on Haystack, Robertson et al. on Scalable Fabric and Task Gallery that use varying scale 2D projection and 3D objects to represent information objects respectively, Voida et al.'s work on extending the 2D desktop surface to wall displays so montages of windows and objects can be continuously and visibly represented, and Kaptelinin and Czerwinski's introductions, all in this volume).

When the number of functions, programs, and files for an average user were relatively small, the desktop metaphor and the point-and-click style of GUI interface had an obvious advantage since users could interact with information objects by visual recognition and reaction, easing the burden of learning and memory. Furthermore, due to de facto standardization, a set of GUI conventions, even some unnatural ones (such as double clicking to open) have become second nature to most users. However, the rapidly growing number of functions, applications, and files (to hundreds if not thousands), puts strain on such the desktop interface, at least in its conventional form. To relieve the strain, desktop search, which enables the user to find files in the local computer without navigating the desktop folder hierarchy, is gaining acceptance. Alternative or extended forms of information representation guided by different metaphors may also gain eventual acceptance. We do not believe that today's GUI conventions can be supplanted by one simple alternative representation having dramatically larger capacity, greater consistency, and the same level of ease-of-entry. More likely in the future a variety of advanced visual representations may be adapted to specific problem domains and different device form factors, complementing the basic conventional desktop metaphor.

Dimension 3: From interaction with one device to interaction with information through many devices

The term "desktop" as computer jargon has multiple, interrelated meanings. One is as the top level "folder" in the hierarchical organization of files and applications in a personal computer. Another is as a set of visual representation conventions loosely guided by the metaphor. But the term also frequently refers to computers that take the form of a "workstation," typically resting on a desk (and by extension, on the lap). Leveraging the economies of scale, this form of computer (commonly known as the personal computer or PC) revolutionized computing from the much less accessible mainframe and timesharing computers. Personal computers give individual users the flexibility of installing and configuring their own software environments. Recall discussions in dimension 1 on information containment in office metaphor, the drawback of relying PCs as the sole information processor is that personal information is trapped in one fixed device (the PC), limiting mobility and flexibility. This is particularly evident for non office workers. See Bardram's observation of the inconvenience of the location and form restriction imposed by today's desktop and laptop computers for doctors and nurses in a hospital (Bardram, this volume).

While desktop and laptop computers will continue to be important platforms of personal computing, non-desktop computers, such as smart handsets, tablets and electronic white boards, will complement today's unipolar desktop personal computers to a far greater extent than today. Consistent with visions of ubiquitous and pervasive computing, all networked digital devices and appliances in many different forms can potentially be connected and hence become interfaces to the personal information cloud. Potentially everyday objects or appliances (Norman, 1998) can also be "powered" by the information cloud. For example an electronic restaurant menu, once opened by a particular individual, can be connected to the individual's information cloud that keeps track of her diet history, preference, and restrictions.

There are many user interface design challenges when the same information can flow in and out of very different devices. How can the same information outflow from different physical devices have enough invariance in appearance and behavior, so that the user can easily identify it and interact with it? How can a unified and logically consistent user experience be provided independent of a device's specific form factor? What can be done to ensure the user has a coherent and consistent human-information interaction experience? For example, a user should be able to interact with his or her calendar events whether the computer at hand is a desktop PC or a smart handset. Separating the data model from its view has long been recognized as an important principle in computing in general and in user interface design in particular (Wiecha, Bennett, Boies, Gould, & Greene, 1990). Initiatives at the W3 consortium in areas such as device independence may lay ground work for achieving transformational user interfaces (Paterno & Santoro, 2003) (Calvary, Coutaz, Thevenin, Limbourg, Bouillon, & Vanderdonckt, 2003); but many difficult challenges call for significant HCI research effort. For example, can a truly usable user interface be designed independent of the specific form factors of a device? How can we counter the arguments that a good UI design has to consider the specific physical form factors of a device? Is there a fundamental set of interaction vocabularies that can be implemented in a variety of device forms so that information can be presented interactively on any device that supports such a set of vocabulary? These issues will be even harder to resolve than hardware independent software development, which has proven very difficult.

Another important topic along the dimension of device diversity is the development of principles, technologies and infrastructure to support teaming multiple devices with different input and output modalities to form a gestalt user experience, so users could opportunistically utilize the advantage of more than one device or information channel to accomplish a task (Ahn & Pierce, 2005) (Yin & Zhai, 2005).

Dimension 4: From mouse and keyboard to a greater set of physical interaction devices and modalities

An integral part of the desktop interaction experience are the physical input devices, in particular the mouse as a pointing device and keyboard as a device for inputting text and evoking commands (e.g., function keys). Almost all software today is designed to rely on these devices. As the personal information cloud model and multiple device form factors begin to evolve, mouse and keyboard can no longer be the only form of physical interaction device. However, the explicit or implicit assumptions of a pointing device and a keyboard are so broadly and deeply adopted in today's software development that even the Windows' Tablet PC, which is quite similar to a traditional desktop and laptop computers in form and size, is markedly more difficult to use. Developing novel, potent yet practical interaction methods that are suited to non-desktop forms of computing is a rare opportunity for the user interface research field. The UI research field in general values novelty, often at the cost of practicality and real world impact. Developing novel yet practical interaction methods is a difficult challenge, since the novel interaction methods are expected to match the performance of the mouse and keyboard, but without the same long learning curve. Experienced computer users have spent years improving their typing and desktop interaction skills, so that even some artificial conventions have become natural to most users. For non-keyboard based input methods to gain favorable acceptance by the users, deep research and careful design have to be invested in developing them. Leveraging users' existing desktop experience and skill, interaction methods that are "transplants" from the conventional desktop may provide a safe path. Paradoxically, such transplants often are poor replications of the desktop experience, inhibiting the full potential of non-desktop computing devices. For example, when using a pen to interact with a point-and-click style of desktop GUI interface, actions that are rather simple for a mouse-based interface, such as a double click, become more awkward, while the dexterity and expressive power of a pen go wasted.

Pen-gesture-based input methods have long attracted both researchers (e.g. Kurtenbach & Buxton, 1994) and product developers (from Go, Apple Newton, Palm Pilot to Windows Tablet PC). Although pen-based interaction methods still have a long way to go before they can truly take advantage of the dexterity of the pen and yet be self-revealing enough to be compelling to novices, many research projects in the user interface field show promise (Hinckley, Baudisch, Ramos, & Guimbretiere, 2005). In our own lab we have been developing interaction models of using pen-crossing action as a counterpart to mouse pointing (Accot & Zhai, 2002) (see also Apitz and Guimbretiére's work on CrossY; Apitz & Guimbretiere, 2004) and a new way of entering text and command using ShapeWriter (also known as SHARK shorthand). Shape writing takes advantage of the fluidity and dexterity of the pen in gesturing patterns, the human's sensitivity in perceiving, remembering and producing geometric patterns, and the modern computing capability in processing statistical constraints to efficiently enter text and commands on non-conventional computers (Zhai & Kristensson, 2003) (Zhai, Kristensson, & Smith, 2005).

As devices become more diverse, the interaction modalities may move beyond pointing, typing or even pen input. Voice and eye-gaze are two modalities that may be taken advantage of in certain situations (Oviatt, 2003). Multimodal interfaces could be particularly effective if contextual information can be drawn from sensing and the personal information cloud, so that these modalities are used cooperatively to their respective advantages.

Progress in the dimension of new input methods faces the challenge of overcoming users' existing mental models, skill sets, and habits. (This also holds for Dimension 2 and perhaps many others.) Making changes concerning the interlock of user skills acquired under a set of conventions tends to be very difficult. Using the QWERTY keyboard as a prime example, Paul David argues a "path dependence" or "lock-in" theory, dubbed qwertynomics, in which an accidental sequence of events may lock technology development into a particular irreversible path (David, 1985). The opponents of qwertynomics argue that the qwerty keyboard has not been replaced because there is no convincingly superior alternative to the QWERTY layout , citing much human factors research (Liebowitz & Margolis, 1990). Regardless of the strength of arguments on either side, innovation concerning user interaction clearly has to either tap users' existing skills and behavior or offer dramatic advantages over conventional practice. Today new forms of computer devices clearly demand alternative input and output methods, but they have to be well researched to be successful.

Dimension 5: Software and computing functions move from applications to services

Today most of the computing functions are delivered through applications residing on the personal computer. An alternative approach is gaining momentum in the computer industry: Server-based computing functions (services) delivered through the internet to a personal device, with internet search being the most successful example. Other examples include web-based email services. There are a number of factors that favor such a shift. First, the trend to being always-connected (e.g. today's push in many cities for a munincipal wifi) enables the viability of the service model. Second, conventional applications have gotten overly complex for most people to make use of or even to know about all of the functions in them. Web-based services tend to be much simpler and "under-featured," perhaps because services can't download huge bundles of code or because these services are young and not yet "enriched." Software services are forced to ask what is really needed, thus enforcing simplicity, which could mean more stable functions. Third, with AJAX (asynchronous java and xml) technologies, web UIs can be very GUI-like, therefore easy to use and familiar in appearance and behavior. Fourth, unlike applications that are difficult to deploy frequently, services can be updated seamlessly (although software service providers really should be considerate of users' familiarity with their interface and refrain from forcing new looks and behaviors on the user every month). Finally, with services users tend to have more choices, since they're easier to find and try out; and potentially users can combine finer-grained services to their individual needs. The shift from applications to services obviously requires a different economic model for business (to date advertising has been the main economic enabler). It also has to overcome privacy and security hurdles.

The shift from application to services also is evolving in parallel with, and faster in pace than, the evolution from personal desktop computing to the personal information cloud model. Together they may significantly influence the form of future integrated digital work environments. Software services should be able to adapt to a variety of individual devices as needed. In a ubiquitous computing world, a variety of devices including desktop computers, handsets, specialized appliances, or in-car-computers could be used to accomplish a task. How could these devices team up effectively in an ad hoc fashion as the user moves around? Applications residing on these devices communicating with each other in a peer-to-peer fashion is a possibility (Newman, Izadi, Edwards, Sedivy, & Smith, 2002). Another possibility is to support a variety of personal or public devices from software services. Based on personal identification sensing or user log in, services in the network could virtually track what devices are being used by an individual, and coordinate these devices and deliver information suited to each of the devices being used. Such a user (ID) centered integration approach has been demonstrated in our FonePal system (Yin & Zhai, 2005), (Yin & Zhai, 2006) in which telephony voice menus are visually displayed on the user's computer screen via instant messaging infrastructure based the user's IDs.

Dimension 6: From personal to interpersonal to group to social interaction

The desktop/office metaphor supports the individual in managing his working set of personal information. But the individual doesn't live in isolation. Although personal information consists of information that is relevant to the person, most of it is not created by the person himself, but by other people. A person's communication with others, such as email or instant messaging, is not only personal, but interpersonal. The metaphor provides an inbox for such communication and also for exchanging information artifacts; but these communications are kept and managed in each person's desktop/office. Interpersonal interaction, by which we mean interactions targeted to specific other people, is not distinguishable from purely personal interaction. The desktop/office can accommodate a range of interpersonal tools. Collaboration (or interaction) with a group or team is where we begin to step outside the desktop/office metaphor. Collaboration is most often supported by some form of "place," such as a "teamroom," where information is shared. What makes such a place separate from the personal desktop/office is that the management of the place is shared with or by others. (Note that here we are not distinguishing how the place is supported architecturally, such as by client-server or peer-to-peer.)

The next level is to engage in more overt social interaction. One aspect of social interaction is to treat people as focal points in the personal information cloud. This is well illustrated by ContactMap and Soylent (Fisher and Nardi, this volume). ContactMap helps a person to explicitly manage his relationships with others, creating a personal social network. To do this we need persistent representations of people and their identities in the personal information cloud. Given people objects, we can organize information around people, such as a history of communications and shared objects. Notice the kinship with Lifestream (Freeman, this volume). Further, as illustr ated in Soylent (Fisher and Nardi, this volume) we can use this same information to infer groupings of people into social and work contexts.

A second aspect of social interaction is making more information (which used to be personal or interpersonal) more available in a wider social context. There seems to be a trend here. More and more services are being created on the Web that encourage people to disclose information publicly. People are putting out information and opinions on personal blogs that are available to an unknown public. People are contributing to various collaborative open source projects, such as the Wikipedia. People are tagging information, such as web pages and documents and photos, and making these tags public to create a system of social tagging for indexing information, often called "folksonomies." Thus more information in the personal information cloud is being made public to combine with others' – creating public information clouds consisting of the intersections of personal information clouds. Perhaps this is a fad, or maybe the web is evolving into a "culture of participation" where public information is created that is greater than the sum of the personal contributions.

Important new social dynamics are emerging, and these must be taken into account, since they will strongly shape the future of integrated digital work environments.

Dimension 7: From low-level tasks to higher-level activities

The desktop/office metaphor provides a set of generic tools for users to work on the information objects in the office. These tools, or applications, support a set of common low-level tasks, such as editing a document, sending an email, organizing a folder, etc. It is up to the user to select tools and use these tools and objects to accomplish higher-level objectives, or *activities*. People think of work in terms of activities (Gonzalez & Mark, 2004), e.g., write a book chapter, and over time perform a series of tasks to carry out the activities, e.g., start a new chapter file, gather related materials in a folder, email the book editor, set a due date in the calendar, edit the chapter, find references in related papers, print the chapter, and so on. The desktop/office metaphor affords great flexibility is organizing the activity, but it offers little help in managing the activity. The activity involves heterogeneous tools and objects scattered throughout the desktop. Many tools do not work well together, e.g., a reference in an email has to be cut and pasted into the chapter file lest it be forgotten.

Many chapters in this book can be seen as striving to support work at the activity level. The Group Bar, the Scalable Fabric, and the Task Gallery (Robertson et al., this volume) attempt to ease the user's ability to manage their activities beyond individual windows and applications. Haystack (Karger, this volume) provides ways to express relationships between disparate objects to organize them better for activities. Lifestreams (Freeman, this volume) replaces the desktop with a stream of document-based actions that can be organized into activities. The notion of roles (Plaisant et al., this volume) can be seen as kinds of activities. UMEA (Kaptelinin and Boardman, this volume) is an explicit activity management system, and their WorkspaceMirror can also be seen that way, as indeed can their general notion of Workspace-Level Design. Kimura (Voida et al., this volume) is explicitly designed to support activities by representing them as montages of document images on a wall display. Finally, the Activity-Based Computing system (Bardram, this volume) develops an explicit architecture and services to support activities in a hospital setting.

The notion of activity is an important concept across the social, behavioral, and management sciences. Most HCI researchers refer to Activity Theory's formulation of activity, e.g. (Nardi, 1996). But there are other relevant perspectives: Distributed Cognition (Hutchins, 1994), linguistics (Clark, 1996), and organizational behavior, which calls them *routines* (Pentland & Feldman, 2005). Activity is also becoming an important analytic construct for understanding usage context in system design (Gay & Hembrooke, 2003; Moran, 2003; Moran, Cozzi, & Farrell, 2005; Nardi, 1996). But more important here is to see that people have to manage their activities and that integrated digital work environments need to support this activity management (Moran, Cozzi, & Farrell, 2005).

Therefore, we agree with Bardram that the activity concept should be made a first-class computational construct that can be used to support human activity. Further, we believe that development of a standard representation of activity, called "unified activity" in (Moran, Cozzi, & Farrell, 2005), could provide a semantic foundation to enable integration across diverse work-support systems. A represented activity is straightforward. Activities are objects with some descriptions (objective, status) related to the people involved, the resources used, and the bounding events. Activities are also related to other activities (such as subactivity). Activity descriptions are fundamentally relational metadata for grouping and organizing elements around human activities (Dragunov, Dietterich, Johnsrude, McLaughlin, Li, & Herlocker, 2005; Kaptelinin, 2003). How do activity descriptions relate to the personal information cloud? Activity descriptions are the part of the personal information cloud that provides organization of that information around the semantics of activity – how the information is used and what it is useful for – the "personal activity cloud."

A standard activity construct can have many benefits. First, it provides objects around which to aggregate the resources to carry out activities, and also suspend and resume activities. Activities are shared information and thus can provide coordination and awareness among collaborators, as illustrated in the Bardram and Voida et al.. chapters, this volume, and also by ActivityExplorer (Muller, Geyer, Brownholtz, Wilcox, & Millen, 2004). Activities are explicit representations that people can operate on, thus providing a focus for reflecting on and planning activities. If activities are represented as they are carried out, then they provide a valuable record of *experience*, which can be reused ("how did George do it last month?"). Another powerful method of reuse is to create *activity patterns*, perhaps by "cleaning up" activity experience records to capture "best practices." It should be noted that activity representations are very different from formal workflow process descriptions in that activities are malleable descriptions under the control of the people using them, and thus adaptable to varying situations. Activity descriptions could complement workflow systems if properly integrated (Moran, Cozzi, & Farrell, 2005).

There are at present only a few research prototypes of activity-support systems (Dragunov, Dietterich, Johnsrude, McLaughlin, Li, & Herlocker, 2005; Kaptelinin, 2003; Moran, Cozzi, & Farrell, 2005) (Bardram, this volume; Voida et al., this volume), and these have raised as many questions as conclusions. There are many challenges to shifting people to an activity-centric mode of working. How are activity descriptions going to be created? Can they be automatically identified from monitoring action streams, as many chapters in this book discuss (Kaptelinin, Voida et al., and Bardram, this volume). It is well known that current automated methods are not accurate enough and require considerable manual "clean up" to make the results useful (Kaptelinin, 2003). Can we do better? Can we create an attractive cost/benefit continuum? It would be extremely easy for users to create crude but useful activity descriptions (e.g. a threaded email conversation could be converted to an initial activity description by a single gesture). Activity descriptions would be further developed because they provide a flexible service for resource sharing, planning, and awareness. Another incentive for using activity descriptions is that can be generated from activity patterns, providing an initial structure and advice. But can we make it easy enough to create useful activity patterns at a useful level of abstraction? And how can we make the patterns available in appropriate contexts? And so on.

Where do we stand?

The theme of this book is that the world is moving beyond the desktop/office metaphor. It is not exactly clear where it is going, but the seven dimensions presented above articulate a design space that is being explored; they chart the course we are on. The diversity of these dimensions suggests that progress will not be uniform along all the dimensions. Research and industry will push forward on different dimensions based on creative insights and commercial opportunities.

We have observed that the desktop metaphor is a caricature of the current state, since we are clearly already well beyond the desktop caricature. So, where do we stand? Let us consider each dimension separately:

1. *Personal Information Cloud.* Personal information has already started dispersing. Many users have their emails, calendar, and documents on the web. However, the shape of a personal information cloud model will take many years to evolve. What is not clear is who will provide the service to maintain and deliver the personal information cloud. The providers could be reputable corporations or open source organizations. Probably there will not be complete end-to-end host providers at all. Rather, the personal information cloud would be organized by a set of services that glue and coordinate their information from multiple hosts and servers.

2. *Diverse representations.* The conventional desktop/office metaphor and GUI continue to dominate, although it is increasingly complemented by desktop search and other new functions. New form factors for information devices are beginning to challenge the status quo and demand alternative forms of information representation.
3. *Device multiplicity*. We already see many forms of computing devices, ranging from handsets to embedded computers in cars on the market. However these devices are largely isolated from each other. Achieving transformational user interface design so that the diverse forms of devices can all be powered by the personal information cloud and deliver much greater value is still at a very early research stage.
4. *New interactions and modalities.* Voice as an interaction modality has finally made practical applications in telephony systems**.** Many other input methods (e.g. telephone pad based input) are alternatives to traditional mouse and keyboards and are already frequently used by mobile users, although existing methods tend to be rather inefficient or even clumsy. User interface innovations in this area have an opportunity to unlock the full potential of mobile and other forms of computing**.**
5. *Software as services*. Software services are rapidly gaining acceptance in the computer industry due to market forces. Already there are available enough services on the web for an individual to do serious work (and most of these are free, at least in limited forms), although some desktop/office functionality is still useful to glue all the services together. This dimension will mostly be led and driven by the intense competition in the information technology industry.
6. *Social interaction.* Social software is surprisingly popular. It is changing the way information is communicated (e.g., blogs), and it is changing the way we think of the web and challenging our assumptions about large-scale social cooperation (e.g., Wikipedia). This dimension is largely based on early research efforts (e.g. Wikis) and is now being driven mostly by innovative experiments and evolutionary progress based on wide adoption.
7. *Activity-centric computing*. The general notion that software should be more activity-centric is widely believed. Current desktop/office environments are slowly evolving in this direction, as are some enterprise collaboration environments. Beyond such incremental changes, there are still only research

explorations, such as those exhibited in this book. There are research challenges in this dimension: the architecture for activity-centric computing, standards for activity representation, and the user experience of being activity-centric vs. being tool-centric and/or inbox-centric. From this research we can expect to see some public experiments and commercial offerings in the near future.

This book presents several research innovations that explore significant steps to the future beyond the desktop/office, as well as the rationale for the directions they represent. We have tried to add some perspective to the work here by laying out seven dimension of change that they are participating in. Although some of the dimensions are strongly driven by the fast pace of commercial innovations on the web, all the dimensions present significant research challenges. Research can guide future integrated digital work environments by articulating human needs and capacities and exploring and evaluating technologies to meet them. The field of human-computer interaction has not had a greater opportunity to impact the broad computing industry, and indeed how people work and live in the world, since the desktop metaphor and graphical user interfaces were invented.

Acknowledgments: We thank our colleagues at IBM Research for creating an intellectually stimulating environment and many discussions that have shaped our thinking on the future of human computer interaction.

References

Accot, J., & Zhai, S. (2002). More than dotting the i's - foundations for crossing-based interfaces. *Proceedings of CHI 2002: ACM Conference on Human Factors in Computing Systems, CHI Letters 4(1)*, 73 - 80.

Ahn, J., & Pierce, J. S. (2005). SEREFE: Serendipitous File Exchange Between Users and Devices. *Proceedings of Mobile HCI*, 39-46.

Apitz, G., & Guimbretiere, F. (2004). CrossY: A crossing based drawing application. *Proceedings of UIST-- the 17th ACM Symposium on User Interface Software and Technology*, 3-12.

Calvary, G., Coutaz, J., Thevenin, D., Limbourg, Q., Bouillon, L., & Vanderdonckt, J. (2003). A Unifying Reference Framework for multi-target user interfaces. *Interacting with Computers, 15*(3), 289-308.

Clark, H. H. (1996). *Using Language*: Cambridge University Press.

David, P. A. (1985). Clio and the Economics of QWERTY. *American Economic Review, 75*, 332-337.

Dragunov, A. N., Dietterich, T. G., Johnsrude, K., McLaughlin, M., Li, L., & Herlocker, J. L. (2005). TaskTracer: A Desktop Environment to Support Multi-tasking Knowledge Workers. *Proceedings of International Conference on Intelligent User Interfaces*, 75-82.

Gay, G., & Hembrooke, H. (2003). *Activity-Centered Design: An Ecological Approach to Designing Smart Tools and Usable Systems*. Cambridge, Mass: MIT Press.

Gonzalez, V., & Mark, G. (2004). Constant constant multitasking craziness: Managing multiple working spheres. *Proceedings of ACM CHI2004 conference on Human factors in computing systems*, 113 - 120.

Hinckley, K., Baudisch, P., Ramos, G., & Guimbretiere, F. (2005). Design and Analysis of Delimiters for Selection-Action Pen Gesture Phrases in Scriboli. *Proceedings of CHI 2005: ACM Conference on Human Factors in Computing Systems*, 451-460.

Hutchins, E. (1994). *Cognition in the wild*: MIT Press.

Jones, W., Phuwanartnurak, A. J., Gill, R., & Bruce, H. (2005). Don't take my folders away!: organizing personal information to getting things done. *Proceedings of ACM CHI2005 Conference on Human Factors in Computing Systems, Extended Abstracts (Short paper)*, 1505 - 1508.

Kaptelinin, V. (2003). UMEA: Translating interaction histories into project contexts. *Proceedings of ACM CHI conference on Human factors in computing systems*, 353 - 360.

Kurtenbach, G., & Buxton, W. (1994). User Learning and Performance with Marking Menus. *Proceedings of CHI: ACM Conference on Human Factors in Computing Systems*, 258-264.

Liebowitz, S. J., & Margolis, S. E. (1990). The Fable of the Keys. *Journal of Law and Economics, XXXIII*.

Moran, T. P. (2003). Activity: Analysis, Design, and Management. In G. C. S. a. S. B. e. Sebastiano Bagnara & Lawrence Erlbaum Inc (Eds.), *Symposium on the Foundations of Interaction Design, Interaction Design Institute, Ivrea, Italy (to appear in Theories and Practice in Interaction Design)*.

Moran, T. P., Cozzi, A., & Farrell, S. P. (2005, December, 2005). Unified Activity Management: Supporting People in eBusiness. *Communications of the ACM*, 67-70.

Muller, M. J., Geyer, W., Brownholtz, B., Wilcox, E., & Millen, D. R. (2004). One-hundred days in an activity-centric collaboration environment based on shared objects. *Proceedings of ACM CHI 2004 Conference on Human Factors in Computing Systems*, 375 - 382.

Nardi, B. A. (Ed.). (1996). *Context and Consciousness: Activity theory and human-computer interaction*: MIT Press.

Newman, M., Izadi, S., Edwards, K., Sedivy, J., & Smith, T. (2002). User interfaces when and where they are needed: an infrastructure for recombinant computing. *Proceedings of ACM Symposium on User Interface Software and Technology*, 171-180.

Norman, D. A. (1998). *The Invisible Computer: Why Good Products Can Fail, the Personal Computer Is So Complex, and Information Appliances Are the Solution*. Cambridge, Mass: MIT Press.

Oviatt, S. (2003). Multimodal interfaces. In J. J. A. Sears (Ed.), *Handbook of Human-Computer Interaction* (pp. 286 - 304).

Paterno, F., & Santoro, C. (2003). A unified method for designing interactive systems adaptable to mobile and stationary platforms. *Interacting with Computers*, *15*(3), 349-366.

Pentland, B. T., & Feldman, M. S. (2005). Organizational routines as a unit of analysis. *Industrial and Corporate Change*, *14*(5), 793-815.

Whittaker, S., & Sidner, C. (1996). Email Overload: Exploring Personal Information Management of Email. *Proceedings of ACM CHI'96 Conference on Human Factors in Computing Systems*, 276-283.

Wiecha, C., Bennett, W., Boies, S., Gould, J., & Greene, S. (1990). ITS: a tool for rapidly developing interactive applications. *ACM Transactions on Information Systems*, *8*(3), 204 - 236.

Yin, M., & Zhai, S. (2005). Dial and see: tackling the voice menu navigation problem with cross-device user experience integration. *Proceedings of UIST 2005 -- 18th ACM Symposium on User Interface Software and Technology*, 187-190.

Yin, M., & Zhai, S. (2006). The benefits of augmenting telephone voice menu navigation with visual browsing and search. *Proceedings of CHI 2006: ACM Conference on Human Factors in Computing Systems*.

Zhai, S., & Kristensson, P.-O. (2003). Shorthand Writing on Stylus Keyboard. *Proceedings of CHI 2003, ACM Conference on Human Factors in Computing Systems, CHI Letters 5(1)*, 97-104.

Zhai, S., Kristensson, P.-O., & Smith, B. A. (2005). In Search of Effective Text Input Interfaces for Off the Desktop Computing. *Interacting with Computers*, *17*(3), 229-250.