Views of work, the foundations of architecture

Austin Henderson
Xerox Corporation, Palo Alto, U.S.A.

Abstract

This paper explores how the perspectives that we adopt in understanding how people work, shape the technologies that we design to be part of that work. Using examples from our current technology, it indicates how the resulting machines, and their user interfaces, fall short in the roles they are designed to play in work. Such descriptions as *user friendly, idiot proof, command, job, work flow* and *supporting work* betray viewpoints which lead to difficulties in our designs. The paper suggests through possible future designs that better understandings of work can lead through the framing of architecture to better technology, systems and working.

1 Introduction

As we engage in designing and developing technology for people to use in their work, we have in mind specific understanding of that work. This understanding may address specific work in a specific situation, or it may be the more generic understanding used when designing for a market segment. It may have come from careful study of work, or from speculation and projection. In any case, the understanding will describe the work either tacitly or explicitly.

Of more interest is the fact that any such understanding will be based not only on the work being addressed, but also on particular views of the nature of work, its characteristics and dimensions. These views of work will be reflected in the description of the work, and more importantly in the technology produced. More importantly than either of these is that the views will be reflected in the work in which the technology plays a part. It therefore becomes either tacitly or explicitly the concern of those working. Therefore the views of work adopted by the designers – and as I will suggest, this will include the workers themselves – are crucial to the design of technology, for they deeply effect the work itself.

At Xerox, I am a member of the Corporate Architecture group. Our job is to help coordinate the design and development not of a single piece of technology, but rather of the whole fleet of Xerox products. Such coordination is necessary because the various products produced often need to work together when they meet in the customer's workplace as part of the customer's work. Further, if these products are based on a coherent view of work, then the customer's understanding can apply

across all participating technology, easing learning and the management of expectations, and simplifying the task of shifting between products because they share certain things in common. (And incidentally, there is the practical advantage internal to Xerox that if products share designs, then they may potentially share implementations and hence production costs.)

However, Corporate Architecture is not directly responsible for designing products; rather we consult with developers in the divisions who do the design. What we do design is the architecture of the products. The architecture is the framing within which, and on top of which, the product design takes place. This framing is precisely the place where the views of what the technology should be are most deeply and consequentially expressed and embedded. By explicitly requiring as part of the architecture that some particular characteristics be addressed in all products, Corporate Architecture can encourage the coordination of designs with its attended advantages.

Therefore, as views of work frame the description of work, the architecture of products frames the description (design) of products designed to take part in that work. It should therefore not be surprising that views of work and tenets of architecture are closely linked. Developers design products to be part of specific certain work. Architects design the framework upon which those designs are based, and those frameworks reflect – often very directly – views of the work. Architects must therefore be deeply aware of the views that they are adopting and building into their architectures (and therefore products and work).

This paper explores some of the views of work that we are building into the architecture of Xerox products. Wherever possible, I take examples from experience with Xerox' products, partially because I know these well, and partially because some of them offer challenges beyond workstation-based software products. This work is as yet quite young, and therefore neither the examples given and nor the points of architecture should be taken as complete or definitive; rather they should be seen as suggestive of the kinds of issues that we are addressing.

The issues are divided into five sections, going from the most general to the most specific. They address: work, the relation of technology to work, the circumstances of working with technology, preparing for working with technology, and working with technology.

2 Work

This section considers work generally, whether or not technology is involved. The issues raised have implications for technology when it is involved, both in that

technology could make it either easier or harder to address these issues, and in that technology could play an explicit role in addressing them.

2.1 Work is social

We start with the observation that people work with others; or even if they work alone, either they work for someone else, or what they produce is used by someone else. This has many results: For one, others are interested in my work: they will be interested in how well it is going, when I will be done, am I doing the thing they think I am. I in turn have to decide which of these requests I will honor, and how I will honor them, and which I will refuse and how. I may set up the work so the requests for status are explicit, or I may arrange things so that others can „look over my shoulder". For example, public copiers are set up in hallways not only so that people can gain access to them easily, but also so that others can see when the machine is free, or that someone has completed a job that you are interested in. In control centers, positions are arranged so that people can easily watch or hear others at work, thereby making the work itself produce as a by-product reports on its status [3, 9]. Other results of work being social are: people talk to each other [11]; people define their work as that of individuals-in-a-team [3]; and so on. A large and growing literature of modern work including anthropological and sociological studies provides insight into people working together.

The implications of the social nature of work for the functionality of any technology are extensive. If visibility of work is important, then designers of technology must be aware when their technology changes the nature of the workspace [14]. For example, Xerox is now offering remote access to some of its machines, machines which have been isolated copiers until now. But if hallway copiers are shared workspaces, then providing electronic remote access to them must be more than providing access to the copying function; it must also include means for seeing the copying that others are doing, and provide ways of communicating with them, ways of passing work off to others as the job evolves [3].

The implications for the architecture of technology with such functionality are equally extensive. In order to „see" work, one must have the concept of „work in action". For example, while Xerox printers permit many „jobs" to be queued at once, remote access to the status of those jobs has not in past been uniformly supported. The architecture should therefore provide the concept of „job" to frame all activity. (We will see later that this concept is not a simple one.) In order to „see" the people who are working, the architecture should provide some concept of person, and the relation between the person and the machine and the jobs they are doing. Interestingly, since local use of hallway copiers does not usually require that people identify themselves, mixed local and remote access may require change in local use.

Or is an unidentified user an acceptable notion? If the architecture does not answer these questions once for all cases (unlikely in this case, since both solutions will be the right one for some worksite), it must provide concepts which can be used to express and implement the answers appropriate to the situation. Similarly, in order to communicate with other workers, communication must be provided somehow, if only in knowing telephone numbers to call.

The implications for design are a change in how we view the work that we are designing for. Simply, we must look for the sociality of work. In a recent visit to a manufacturing facility where I had been asked to comment on the user interfaces of the machine at use there, I was shown around for a few hours, at the end of which time I asked my hosts whether there was much multi-person use of the machines - that is, two or more people using the machine at once. They had difficulty answering, not I think because they had not seen such activity (over half the cases I saw of people using the machines involved two people), but rather they had not been orienting toward that dimension of the work. I might have gotten the same answer had I asked them how many workers were wearing red clothing, or were left-handed, or a thousand other things. I had been looking for the sociality of the work, and could supply an answer; I have no idea about red clothing, and only a guess at handedness (which is of concern in user interfaces too).

Two questions arise: Which of the thousand things that one might look at in examining work should one address? The answer turns on what aspects of working have implications for functionality and architecture. And: how does one look in order to see and hence to know what matters? Those who study people have ways that we can learn from: anthropology and sociology have some powerful methods (see, for example, [2, 10]), as do the design sciences [6].

I have taken some space here to touch only very lightly on one dimension of work, that it is social. While this coverage does severe disservice to the understanding that must be achieved to properly address this reality, it must suffice for this paper, as it is only one of many issues. The discussion does, however, exemplify the relationships between facts about work, and hence the ways that we must view it in order to understand what we are doing as designers, implied functionality and architecture, and some possibilities for design and the effects on people working. In the rest of this paper, I will not have space for as much detail on all the issues. Instead, I will leave it to the reader to fill in those aspects I omit.

2.2 Work is a tension between regularity and difference

Work is forever new. Each activity is a new activity with its own particular circumstances. To be true to the demands of those circumstances, workers must respond to

each activity particularly, producing responses appropriate for the particular case. There are many resources available for handling each activity, most powerful among them being experience with previous activity which is in some way like this activity. That is, work is also old, in that an activity may be regarded as being „the same as" previously encountered activity. Over time, patterns of sameness may be noticed and these become the regularities that are oriented toward in each new case. Some of the „work" of work is then seeing new activities as being cases of established regularities. There will of course be specific circumstances of the case in hand that will be unique to this case, and not regarded as part of the regularity.

For example, Xerox sells supplies for Xerox copiers (paper, toner, etc.); customers can call a supplies telephone number and arrange for supplies to be delivered to them. The worker who answers the telephone sees their work as having regularities; specifically, they see their work as including the „taking of orders". This is only one of the regularities that an order-taker faces; others include providing information about supplies, and giving status of previous orders, answering requests from other order takers, talking to their manager, and – when their children call – answering when they will be home that evening. However, having recognized a call as being an order-call, the worker can then use the details of the regularity of order-taking as a frame for the activity. The „work " of order-taking then continues with matching those details against the realities of the particular situation. So for example, a name is required, and an address. The regularity does not specify the name, or the address; that is specific to each case.

However, reality being what it is, the regularity can prove to be unsuited to the particular case in some way. For example, when asked for an address, one customer whose copier was on an ocean-going barge could only ask in return when the supplies were to be shipped, because the address for shipping would vary with time. The order-taker got a telephone number, and supplied it as an address to the shipping department, with a note to, „Ask for Bob", when the order was ready to be sent. The order did not quite match the regularity, but the order was completed nevertheless. Further, the experience of order-taking was enlarged as a result of this encounter with a new reality: The notion of an address was extended to cover the idea that copiers move, and the notion of recording an address was extended to cover providing a procedure to follow to determine the address when the address was needed. If this happened often, the regularity of order-taking might be extended to encompass this contingency. However, reality is infinitely rich, and the world is producing new situations all the time. Not all experiences will be recognized as cases of regularities; in fact most details are not regular at all.

We can view work then as produced by, and producing, regularities, and the work itself being a tension between seeing things as regular and see them as „just" a fact of

the situation at hand. Note that this view of work is not a function of technology; people will see the world as more or less regular, as instances of patterns, whether or not technology is involved. And of course this is compounded by the sociality of work: if others have to agree upon and use the regularity as well (e.g. the shipping clerk had to know what to do with the telephone number) there will be more explicit work necessary to change the regularity.

In passing, it should be noted that institutions, both social and technical, are one of the primary sources of regularities that workers must accommodate, and, in accommodating, re-produce, possibly-altered. The order-taking was set in a situation of companies, contracts, agents, telephoning and shipping; the workers were acting with the confines of gender, race, religion, personal needs, and cultures. Although some of these may seem unrelated, upon reflection one can see that these institutions frame and direct action. For example, the institution of selling brings the pressure to reach a solution to the problem of supplying an address: it is primarily the responsibility of Xerox, the seller, to provide the solution (by foregoing getting the address, and making arrangements for getting it later), not the customer, although the customer is expected to cooperate (by agreeing to answer the telephone and provide an address later). The study of institutions in work is of great interest to us at Xerox because they are such profound providers of regularities in the form of norms.

The functional implication is, again, simple and profound: people working are constantly creating and adjusting regularities in their work to make sense of the world they are in. This means that the natural concepts of the work will be, in the absence of strong forces, constantly shifting. And the architecture? One view is that it should recognize and support constantly changing concepts. An alternative view – one more likely to be realized in today's technology, and more acceptable to the auditors – is that people will map between the world and whatever systems they use in their work; people will use the systems where they can, giving that use meaning in the world through the mapping. Under this view, the architecture should enforce the concept that all concepts are subject to annotation [7]. This can provide at least the mechanism for remembering and communicating the mappings that people have made to the world. The telephone number and the note to „Call Bob" were just such an annotation on a system (paper forms, pencils, and humans) which expected a shipping address.

I am not sure that annotation is the best, or even a good, solution; but whatever is chosen it must deal with the reality of ever-shifting regularities in work. Of particular interest in any design situation is to watch for the forces that drive change, and also for the forces that keep things the same (see below); we must understand what can (and will) change and what will not. Viewed this way, work becomes richer but more understandable, and the place of systems and technology more manageable.

2.3 Work is multi-roled and multi-tasked

Workers have many roles. An order-taker might in addition be a member of the order-taking staff, an employee of Xerox, a member of a task force on new technology, a user of a particular health plan, a representative on the cafeteria committee, and an individual. These different roles require different things of workers at different times, and at the same time. Their preferences, rights for access, indeed their very identity may change with their role. To recognize this, the architecture could provide concepts like role, and extend the registration with the system to indicate what role one is taking. Further, the relationships between roles could be enriched, so that an administrative assistance could log in as themselves, but still have access to the their boss' mail, but not his personal files.

Similarly, workers, even in a single role, are often working on many things at once. While talking to a customer in taking one order, an order-taker may be finishing up details on another order, and handling interrupts from fellow workers. While copying one document, a secretary may be packaging up the copies of the previous document, removing the staple from the next document to be copied, talking to the person waiting for the copier when they are finished, and reassuring their boss that the minutes will be ready in time for the meeting. The need for multi-tasking seems ubiquitous in work, and in designing systems we must recognize it. Workers must be able to shift quickly between jobs, and keep tack of what is going on; when they forget where they were they must be able to work it out from the work; and they must be able to do this without undue overhead. The demands on architecture are that these swapping tasks be recognized as part of the work. Concepts like „task" or „engaged tool", and mechanisms like „Rooms" [4, 8] can provided system support. The demand on designers is that this aspect of work be recognized and observed, and its implications considered in the design of systems to support work.

2.4 Work is described

There are many reasons to describe work: to get funds for it, to justify its continuation, design systems to use in doing it, or even to „automate" it. Of course, all descriptions are partial, and by virtue of that regularizing: by saying only certain things about work, listeners may assume that one has said all. And if all is described in any finite way at all, the work has been reduced to whatever is described. Untrained listeners may also hear the regularities presented in a description will be all that occasions activity in any situation. These two tendencies lead to two important difficulties when thinking about work, for any of the above reasons. First, designers tend to design for the work the description implies – the purely regular work. And

second, listeners tend to think of work as regular, as witness the notion of „work flow" that has been popular for 15 years at least.

Instead, we must insure that we understand that descriptions, including the regularities that workers use, are indeed partial and are of course subject to interpretation in any given situation of work in the world. As architects we must look for mechanisms for putting these descriptions, partial though they are, to good use. In particular, if systems behave as described, they will not be mapped to the world as it is. If workers are given the ability to bring their own judgment to the activity of such systems, to guide and direct that activity as it is happening, descriptions – limited though they are – can be put to good use. (See section 6 on working and rich control.)

2.5 Work is located

A simple view is that people work where they are. However the social view of work means that work can be with someone else or dependent upon something someone else is doing. Also, work can be spread around a large room, or throughout a building, or – with technology – around the world. And putting the two together, people may not even know where their work is. Thus people do not need to be near their work; rather, they need to have ways of finding where it is or what it is dependent upon, of accessing it, of controlling it or communicating with the people who have or are collaborating on it. Architecturally, this requires concepts of units of work, people/roles, dependency, location. For designers it means seeing these concepts being worked out in the work we study.

3 Technology in work

I now deal with issues arising from technology taking part in work. Although I regard technology as including paper and pencil, physical artifacts and space, here I will focus on computational technology: computers (e.g. terminals and mainframes, workstations, personal computers, personal organizers) and computer-based machinery (e.g. copiers, washing machines, watches).

3.1 Work is co-defined by technology

We sometimes talk as if we can define work independent of the technology involved in it. This is most clear when we discuss analyzing the work with the idea that we

can then „do it differently using" a different technology. This idea is also captured in the idea of „supporting" work with technology. Instead, when the technology is changed the work is changed; the „it" that would be supported is a different it. For example, when the control panels on copiers were changed so that information was presented electronically rather than with paper „flip cards," the work of getting information change significantly: the information seen was that provided by the copier, not that chosen by the user which was often faster and easier, but was sometimes not what the user wanted; it was hard to flip between kinds of information; and the ability to write on the flip cards (e.g. to give the name of the person responsible for the copier) was lost. In ways the work was easier, and in ways harder; but the best way to think of it, is that the work was *different.* Other evidence that work and technology are mutually-defining appears in the descriptions that workers give of work: the descriptions tend to be phrased in terms of the technology, even when only discussing the work. For example, people refer to the work of getting information as „looking at the flip cards" or „checking the instructions on the display". And conversely, technology is often described in terms of the work: for example, the count of copies will be referred to as the „number of sets of meeting minutes" and the paper in tray 2 will be referred to as „the covers". To suggest this interdependency, I refer to technology as being „part of" work.

As designers we must therefore think, not in terms of designing a technology, or designing work, but rather in terms of designing the pair. We can learn from existing pairs, but the introduction of new technology will surely change the work. For architects the implication is more profound: if we include in the technology descriptions of work in which it is involved, we may expect those descriptions to be phrased in terms of the technology at hand. That is, the fact that users are already thinking about their work in terms which include the technology means that the technology may reflect itself in its descriptions too. Thus descriptions of technology made for technological reasons may also serve in the description of work. On the other hand, descriptions of technology are better, unfortunately, if they are given in terms of the work being done; this is not known until the work is in progress. However, for work that is done regularly, such descriptions can be created, and presented if the worker indicates the kind of work being done, indications which could come from the user's choice of pre-stored settings and tasks.

3.2 Work is regularized by technology

Because technology is often hard to change, it is common for people to fit the work to the technology. That is, the work will be driven by what the technology can do. Because technology, separate from the people working with it, is not aware of what

is going on around it, or even the meaning of what it is doing, it is very regular. When computers were introduced into order-entry, for example, the technology had built into it a fixed notion of an order, and every order – at least as far as the technology knew – met that description. In mapping the world into regular technology, workers can adapt the regularity of the technology to the richness of the world. However, to do so requires effort, and not only in creating the mapping, but also in carrying it through. So, the unanticipated time-dependent address required that the order-taker create a way of putting that information into the system, and also that the shipper notice the strangeness, understand the information there, and interact with the customer. Because such adaptations of the technology are work, there is a natural tendency for people to avoid adapting if they can. As a result, there is a pressure from the regularity of technology to make the work regular, and therefore for the workers to want to see the world as regular. Thus, the order-taker might prefer to guess at a date, and settle on an address rather than adapt the system. This pressure is increased when order-takers are regarded as interchangeable, and so may not be able to insure that they will be present to carry through on the adaptation.

A further pressure is inadvertently exerted by those who create the technology when they slip from trying to do the best job possible to make the technology match the world to believing that the technology succeeds in matching the world. For example, some system analysts tend to think of themselves analyzing the world so that they can make technology exactly match it. When technologists ignore the fact that workers adapt technology to the world, they are likely to make the technology incapable of helping workers in this adapting process. For example, the order-taking software required that an address have a particular form, thus making it difficult to put in the telephone number and the note. Smarter versions of the program might be even worse, in that they would do a better job of enforcing something that was not wanted – at least in this particular case. The worst result is that the technology defines the work: the instrumental becomes definitional. In these cases workers simply have „work around" the technology.

The architectural implications are that technology should reflect the fact that work being done (e.g. the information it carries) is being mapped by workers into the world. This can mean either that the information in the technology is enriched to reflect the mapping, or that access to that information is enriched so that people can monitor and adjust it when it is relevant. Some designs for enriched information: Have a single value that means „this is an exception, see a person" which can be used in any data field; add information about which person to see [5]; make it possible for any value to be annotated. Some designs for enriched access to information are discussed in section 6.

3.3 Work is structured by technology

Technology brings new structures to work, structures that limit and define the work. These structures arise from properties of the technology: for example, technology is physical and therefore has location; technology has limited capacity and therefore may have to be scheduled; technology is owned and therefore has differential access; technology is expensive, and therefore usage has to be paid for; technology requires specialized knowledge, and therefore is not universally available. Each of the properties is reflected back into work in which the technology takes part: Hallway copiers and printers require walking down the hall; copiers can be used by one person at a time, and therefore work involving copying may require negotiating with others over who may use it next; ownership plays a role in these negotiations; payment may limit the amount of work that can be done; special knowledge restricts who can do it. People learn these structures, and modify their work accordingly. People delay and collect their copying tasks so that they only have to walk down the hall, negotiate, and wait once. Therefore the work that people do with technology is often grouped or sequenced, assigned to workers, not by the forces of the work, but by the forces of the technology.

The architecture should therefore reflect the fact that work is being done at times, in places, by workers dictated at least partially by the technology of that work. If technology changes, the structure of the work changes: for example, if a copier is unavailable, broken, out of staples, out of yellow paper, those affected need to know; if a machine moves, people need to know. The technology can be carrier of this information, a way of helping people get the information they need to structure their work and to coordinate it with others.

3.4 Work is increased by technology

When technology becomes part of work, the work – as a whole – must be extended to include choosing, funding, acquiring, owning or leasing, monitoring, and maintaining the technology itself. Even if this work is done by others (e.g. management, purchasing, maintenance), those who work with the technology must interact with those others. Thus work is extended by technology.

Architecturally, again we see the need for the technology to carry the information which allows people to communicate with others: e.g. for workers to know who maintains the machine. In addition, since the technology is itself the subject of these communications, technology can keep an eye on itself by logging its own activity, monitoring its own workings, tracking its own configuration.

4 Circumstances of working with technology

This section addresses the circumstances under which people work with technology. These circumstances are key drivers of the architecture of technology, particularly the architecture of the user interface.

4.1 Workers access services in different ways

Different workers, because of their circumstances, will want to have different access to any particular functionality provided by technology. A Canadian in Quebec will want to use French, the metric system, Eastern (North American) time zone times, and Canadian currency. If they have a Macintosh, they will want the style of user interface to the service to be in the Macintosh style; if they are standing beside the printer, they will want to use the buttons and lights (hard panel) on the machine. Maybe they prefer a direct manipulation style; maybe pulldown menus. Indeed they may not even want to access the functionality directly, but only as part of another functionality (e.g. accessing scanning, not directly, but as a part of copying).

It is therefore required of the architecture to separate the functionality from the access to it. The functionality is grouped together into *services*. The access to services for workers is through *user interfaces* . Both user interfaces and other services access services through a common *service protocol boundary*. A further consequence of wanting to access services through other services (that is, the desire to composed services) is that services should do all their interactions with workers through their user interfaces, that is to say, through their service protocol boundaries. Then, a composed service can intercept or pass on interactions. Thus services think of delivering service, not interacting with users.

Another consequence of letting the user interface map the service to the circumstances of the user is that all interaction through the service protocol boundary must be in language-independent form. It is the job of the user interface to then express the service concepts in terms appropriate for the local circumstances (e.g. French, metric, Eastern time zone). This is similar to the way that current applications, programs or operating system use resources which are language-specific [1].

4.2 Workers have to access services

Both services and user interfaces are of benefit to the user. With access to services provided through user interfaces, workers need not be concerned with what machines those services run on. For example, in distributed systems (e.g. file

servers on networks), the exact whereabouts of the service may never be known to the worker. Further, as a user interface may run on a different machine than the one actually making the displays (e.g. the user interface is an X client running on a machine remote from that running the X server which makes the displays), this disregard for location of machine extends to user interfaces as well.

Architecturally, this means that we make a strong separation between the „platform" perspective which describes the technology as machines and software, and the „functional" perspective, which describes the technology as the user conceives it (the user's conceptual model). This separation provides for services and user interfaces to span machines, and while conceptually separated to share machines.

However, this separation of machines and function must align at the point that the worker needs to make physical contact with the function. Such contact happens when a service affects the physical world (e.g. a printer produces paper), a user interface presents controls to the user (e.g. lights on a printer, windows and widgets on a display), or a machine which is running services or user interfaces is in some way damaged or unavailable (e.g. the air-conditioning is being fixed in the file server room). In all these cases, the user must know at least the physical location of these particular machines.

The architectural requirement is therefore that physical machines must appear in the functional perspective. We use the term „device" for functional notion of physical machines. Of particular interest to those involved with user interfaces are the devices which make user interface presentations. We call these „portals" to suggest that these are the worker's way into accessing the services.

Thus on one hand, user interfaces make presentations on portals, and on the other they access services through the service protocol boundary. Their job is to *mediate* between the worker and the service, providing appropriate conversion and matching as required.

4.3 Workers do many things at once

Workers may be engaged in many things at once. Furthermore, each of those tasks may require the functionalities of a number of services, some of those services may be required in a number of the tasks [4,8]. For this reason, the architecture must enable a portal to provide for a means of presenting a number of different user interfaces simultaneously (or, under control, sequentially) and also the ability to focus on those user interfaces associated with a particular task.

The architecture must therefore introduce some notion of a workspace of user interfaces, and some workspace management system (e.g. Rooms). However, rather than introduce these as new notions, it is possible to regard these merely as yet more services with their own user interfaces; these service manipulate other user interfaces.

4.4 Workers move around

Work demands that workers move around while doing their work, to access technology, or to discuss the work with others. In future, portals may be portable, as suggested by some of the work on ubiquitous computing [15]. If, however, portals are fixed (e.g. today's workstations), then access to the same service through multiple portals must be provided. For example, I may print a document from my workstation, and then walk to the printer to pick up the hardcopy; if no paper has appeared when I get to the printer, I must either walk back to my desk or access my print job using a portal at the printer. Indeed, portable portals will not replace the fixed ones, but only provide additional possibilities for how to work.

The requirement for the architecture is that these multiple user interfaces remain consistent with one another, even if they are being delivered on quite different sorts of portal (e.g. a Sun workstation under Open Look, a single small window on a handheld device, and lights and buttons on the printer). The state of the user interfaces must therefore be shared between interfaces, and reflected coherently on all of them. Some of this work is easy, in that the state of the user interface is partially a reflection of the state of the service session, and so all interfaces will be able to get it from their service. But if there is additional state in the user interfaces (as in general there will be: e.g. layout details, which section of a complex interface is being viewed, which error message of the five active is being displayed in the two-line seven-segment display) then this state too must be shared and re-interpreted at different user interfaces. An architecture for this is as yet unsettled; we are working on it.

5 Preparing for working with technology

In focusing on working with technology, it is easy to concentrate on a worker engaged with technology doing their work. What this misses – of the work that the worker must do, of what the architecture must address, and of what designers must consider – is the preparation that must be completed before the worker can begin working. Specifically, the workers must manage the situation, they must learn about the situation, the technology and the work, and they must customize the technology

to suit the situation. In fact, these „overhead" tasks are not restricted to coming before the work. At least as often they come in the midst of the work, when some additional management, learning, or customizing must be engaged in. Indeed they are often so intertwined with working that to see them as separate is purely an analytic convenience. Also, these activities are not always required; they are either optional – engaged in to improve the work, or contextual – handled by others as part of establishing a workplace that works. The descriptions of these sections are much shorter than they deserve, mostly because these sections are closer to what we all know as human-computer interaction and are therefore well covered in our literature.

5.1 Preparing the context: managing

Since work is social, and technology may be used by many people, workers must manage their access to the technology. Further, since technology defines and provides resources, and as these are limited, worker's use of resources must be managed. Some of this management is implemented by the physical space surrounding a machine: If there is only a single control panel, and it can only be used by one worker, then managing access to the machine can, and is, left to the social mechanisms that people employ standing in front of the machine. This makes the requirement for this management work hard to see, precisely because it is so commonplace that we find it unremarkable and fail to even take much notice of it. However, when control is distributed out to workers' offices, management will have to be provided through more explicit means: calling on the telephone, registering with the machine, and so forth.

The architectural concepts necessary to address management are those dealing with who people are, what resources are (authentication), what resources people have rights to (authorization), and what their units of service are, how do you know when they have been invoked and completed, and how are they to be paid for (accounting). Copiers have had an extremely rudimentary version of these concepts in past: access has been controlled physically in many machines, and with „auditrons" in others. In future, particularly with the removal of alignment between services and machines, this will need to be much stronger. At the same time, the simple cases of simply walking up and copying cannot be encumbered with logging in unless there is a very easy way of doing it.

5.2 Preparing the worker: learning

One of the clearest lessons from work on operating copiers has been that a major part of the work is learning what the machine does, how to map your work into those

capabilities, and how to request the actions you want [13]. The task of operating a copier is vastly different when you understand these things and when you don't. Of course, the fact that learning is necessary is understood, and various mechanisms have been put in place to address it. However, the most powerful mechanism, the one most often used, is often overlooked because it is not technological: People turn for help to other people. It should be noted that workers must learn about all aspects of work, of which technology is only a part. And since work is locally defined, it is necessary that mechanisms are put in place to help with learning must be extendible to include information about the local circumstances of the work.

The architectural concepts for learning are again related to the objects that make up the technology. There should be ways of referring to those objects, annotating them, finding out who knows about them and who uses them. These concepts in turn require notions of people and locations, in addition to the parts of the technology – service, user interface and device.

5.3 Preparing the technology: changing

Most technology has installation dialogs. Workers have preferences expressed in interaction with applications, or in files that are accessed by those applications. „Design spaces" such as spread-sheets provide workers domain-specific development environments for creating functionality of their own choosing within the target domain. And „construction set" environments (e.g., the UNIX shell scripting mechanisms in which primitives are put together with computational languages) enable workers who recognize patterns of usage of technology, to compose requests for those usages into new services to meet specialized needs. These mechanisms are powerful ways of changing technology to suit the needs at hand. However, they will only be used if it is socially and organizationally acceptable to do so.

The architectural implications of changing the technology are profound. They require that the technology be visible, that the model of its behavior be understandable (in a society which barely understands mechanical devices, understanding computational mechanisms may be hard), and that the resulting configuration be determinable so that maintenance activities know what they are attempting to maintain. Also means for coordinating changes with others who are involved, communicating about the changes, sharing changes and changed technology so that others can use it are all central to enabling workers to adjust technology to meet their needs. Key architectural concepts include configuration, dependency, and specifications of functionality. And these new technologies must be learned, managed and further changed.

6 Working with technology

So finally we come to actually getting something done with technology. Here, perhaps as much as anywhere, the view we take of these activities determines what the technological architecture must be. There are many issues here, but I will mention three that I think are particularly important, all of which are part of a single general theme: Working with technology is best thought of, not as commanding technology, but rather as working together with it to get your work done.

6.1 Working defines the work: figuring it out as you go along

Workers start out with a rough idea of what they want to do. As the work of doing it develops, various realities come to light: Some were predicable if you'd only known; others result from changing circumstances. As a result, the work gets redefined, if only because it is understood in more detail. Sometimes it is necessary to change the original definition of what you were trying to do, and to adjust the work accordingly. For example, in our copier studies, one pair of operators wanted to make 35 copies of a document for distribution. After working on achieving this for half an hour they had refined the task into getting one copy done correctly to learn how the copier worked, and then making 34 more. They had succeeded in the first part and had launched the machine off to make the 34 additional copies. After two copies had come out they could see that it was going to take much longer than they had to complete all 35. So they re-defined their aim as getting out ten copies that they needed immediately, and 25 more to be done by an assistant in the following few days. They then had to redirect the machine to stop at ten, and make the request of the assistant. This case is not at all unusual of „routine" work. Incidentally, notice that the concept of the „job" becomes harder to define under this view of redirecting, making new requests, making multiple requests. A much richer taxonomy of working is needed.

The implications for architecture are that the activity of doing the work is an important part of defining the work, and so the understanding of the work is very likely to change during its course. One key aspect of enabling this behavior is that the technology enables the worker to see what is happening as it works. In this case, the time taken for each set was visible, and so determining that time was inadequate could be made. On the newest Xerox production publisher, copies are made so fast that the tray for completed sets is protected with a cover. This cover lets one see that copies are being made, but it does not permit seeing the copy quality, something that many workers monitor as the job proceeds. The new cover interferes with continuing engagement with the work.

6.2 Interrupting the technology: stopping things in their tracks

Workers may want to stop on-going activity, either to evaluate it, or to change the direction. In the case above, after attempting to change the requested count without stopping the machine, the workers tried to stop it. Their immediate problem was that there are two stop buttons, one which stops the machine immediately, and the other which stops the machine at the end of the next set. The problem, in addition to learning this, was that the immediate stop does not permit continuing, and the „end of set" stop also puts the machine into an interrupt mode. Neither of these was what was wanted. More generally, the issue of specifying how to stop is a job in itself.

The architectural implication is that we need a set of concepts for talking about interruptions. Points of interruption must be defined, and what activity can continue after the interrupts. While it may be possible to get solutions for specific services (as Xerox has done over the years with print and copy), a much more general solution is needed, which works well with services of all kinds, even the composed ones created by workers themselves. In fact, interruption must be part of the description of services.

6.3 Changing the request: setting new directions

Finally, workers will want to change the specification of the activity in which the technology is engaged. After stopping the machine, the workers in our case discovered that they could not change the requested copy count. In the end they simply had to wait it out and stop the machine by hand at the end of the tenth set. But even supposing that the service had allowed them to change the copy count, what would have happened had they requested 2 copies, after 3 had already been made? In general, what aspects of the request can be changed, when? How do new requests fit with existing (and possibly on-going) action?

Again, we need a much richer understanding and set of concepts for redefining activity. Included in this set of ideas must be the objectification of the activity so that it can be referred to and manipulated [12]. It will require forms of „undo " that permit going back to an earlier point, changing things and proceeding „from there." This must also be understood in the context of the machines taking action in the physical world. While it is easy to think about backing up to an earlier point, implementing may be much harder. While it may be possible to back computations up (but what about side effects?), it is hard to „unprint paper". We are working on these issues but have no general answers yet.

7 Conclusion

In this paper I have focused on the implications for architecture of the views that we take of work. From work in general through the fine-grain details of working with technology, I have selected representative issues which indicate how broader views of work and working with technology place strong requirements on the architecture of the technology that we build.

It will be noted that I have not focused on any particular work at all. In particular, the work that Xerox – The Document Company – concentrates on appears only incidentally in examples. Yet the use of documents in work is a matter for serious study, and the understanding of that specific work also leads to points of architecture; for example, the architecture of documents, of printing, of distribution are all concerns of Xerox's Corporate Architecture have not dealt with these matters here in favor of addressing matters which I believe are more fundamental and therefore are pertinent to work of all kinds.

By watching carefully the work that people do, not with the eyes of the technology creator, or even the system builder, but rather with the eyes of the worker, many „obvious" concerns not currently addressed by systems, technology or their architectures are readily discernible. That workers must address these concerns presents significant opportunity for improving the technology that is available for doing work. Possibly even more importantly, by understanding the nature of work and the role that technology plays in it, we can better set the expectations of all concerned for how work is done with technology. The contribution of architecture to this endeavor is to focus attention on these concerns by insuring that mechanisms are available within the technology that permit workers (and other developers) to incorporate it into systems for doing their work.

8 Acknowledgments

The ideas reported in this paper reflect the thinking of the members of the Corporate Architecture group of Xerox Corporation. In particular, I wish to acknowledge the collaboration of Annette Adler on the social perspective on architecture, and the managerial support of Ted Smith and Ron Rider. Corporate Architecture also borrows freely from the thinking of others in Xerox. I wish particularly to acknowledge the work of the groups lead by Paula Siviy in a development division, David Crocker in advanced development, Lucy Suchman in the research laboratories, and Tom Herceg in industrial design.

9 References

[1] Microsoft Word user's guide. Microsoft Corporation, Redmond, WA, 1991.

[2] Atkinson, J.M. and Heritage, J.C., Ed., Structures of social action: studies in conversational analysis. Cambridge University Press, Cambridge, 1984.

[3] Bentley, R. et al., Ethnographically-informed systems design for air traffic control. In CSCW 92 (1992. Toronto, Canada). ACM, New York, pp. 123-129.

[4] Card, S.K. and Henderson, D.A., Jr., A multiple, virtual-workspace interface to support user task switching. In: CHI+GI, 87: Human factors in computing systems (1987. Toronto, Canada). ACM, New York, pp. 53-59.

[5] Fikes, R.E. and Henderson, D.A., On supporting the use of procedures in office work. In: First Annual National Conference on Artificial Intelligence (1980. Stanford, CA). American Association of Artificial Intelligence.

[6] Greenbaum, J. and Kyng, M., Ed., Design at work: approaches to collaborative design. Lawrence Erlbaum Associates, Hillsdale, NJ, 1991.

[7] Henderson, D.A., Jr., „Escape Clauses": a solution for the problem of bounded technology in an unbounded world. In: Ninth Congress of the International Ergonomics Association (1985. Bournemouth, UK). IEA.

[8] Henderson, D.A., Jr. and Card, S.K. Rooms: the use of multiple virtual workspaces to reduce space contention in a window-based graphical user interface. ACM Transactions on graphics, 5, 3 (1986), 211-243.

[9] Hughes, J.A., Randall, D. and Shapiro, D., Faltering from Ethnography to Design. In: CSCW 92 (1992. Toronto, Canada). ACM, New York, pp. 115-122.

[10] Jordan, B. and Henderson, A., Interaction Analysis: Foundations and practice. The journal of the learning sciences, , (1993), (in press).

[11] Minneman, S.L., The social construction of a technical reality: Empirical studies of group engineering practice. Stanford Univserity, 1991.

[12] Smith, B.C., Reflection and semantics in LISP. In: 11th Annual ACM Symposium on Principles of Programming Languages (1984. Salt Lake City, Utah). ACM, pp. 23-35.

[13] Suchman, L., Plans and situated actions: the problem of human/machine communication. Cambridge University Press, Cambridge, UK, 1987.

[14] Suchman, L., Constituting Shared Workspaces. In: Communication and Cognition at Work, Y. Engestrom and D. Middletom, Ed., Cambridge University Press, 1993, (in press).

[15] Weiser, M., The computer for the 21st century. Scientific American, 265, 3 (1991), 94-105.

Austin Henderson
Xerox Corporation
Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, CA 94304
U.S.A.

e-mail : henderson.parc@xerox.com or henderson.chi@xerox.com