Sharing and Browsing Media on a Digital Tabletop

Ali Mazalek
Assistant Professor
Synaesthetic Media Lab, Georgia Tech, Atlanta GA
mazalek@gatech.edu

Glorianna Davenport
Principal Research Associate
Media Fabrics Group, MIT Media Lab, Cambridge MA
gid@media.mit.edu

Matthew Reynolds
Co-Founder
ThingMagic, Cambridge MA
matt@thingmagic.com

Sharing and Browsing Media on a Digital Tabletop

Abstract

As consumer digital media technologies evolve, there is a need for new kinds of platforms that support sociable interactions to manage and display our ever larger personal media content archives. While today's desktop PCs provide a range of applications for media management and sharing, their interfaces force a separation between digital interactions with stored media content, and the physical "daily living" spaces of our everyday lives. This generally precludes the casual face-to-face interactions that have traditionally emerged during the browsing of physical media artifacts such as family photo albums.

In this paper we present two applications for managing and browsing personal media collections on the TViews table, an interactive tabletop display platform that supports multi-user interactions through real-time tracking of tagged objects. The TViews table is designed for everyday environments and supports casual and sociable interactions around media collections, thus removing the separation between physical and digital interaction spaces in the home.

Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentation]: User Interfaces---*input devices and strategies, interaction styles*; H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems--*image/video retrieval*; H.5.3 [Information Interfaces and Presentation]: Group and Organization Interfaces---*collaborative computing*

Keywords

Media table, personal media collection, tangible interaction, sharing, browsing, social interaction.

In recent years, the widespread appropriation of consumer-end digital media technologies such as video and still-image cameras has resulted in ever-larger personal media collections that serve as growing records of our day-to-day lived experiences. Simply put, we collect more and more personal media in the form of digital videos and images, filling endless amounts of storage space and bringing up questions as to how we might access and make use of this vast repository of personal memories. As the technologies to support this growth in personal media collections continue to improve and evolve, so too do the applications and interfaces that serve to access and share them via existing computing platforms, such as desktop PCs and personal mobile devices. Yet there has been little progress in recent years at the level of new interface approaches for jointly sharing and browsing all of these media archives in a sociable manner. The interactions we have with computers and mobile devices all happen one-on-one between the person and machine, or between many people connected by machines, and there is little or no face-to-face interaction that takes place. Designed for workplace efficiency and single-user task-oriented applications, these interfaces have long dominated the realm of human-computer interactions. As a result, many media-related leisure activities which were formerly shared between small groups of co-located people have been retro-fit onto interfaces that were specifically designed for single-user interactions.

In the past, media content browsing was typically not a single-user activity. Consider for instance the way we browse pictures and tell stories. In the days of physical photographs, this was a shared physical activity that happened amongst friends and family members, often in the living room where people gathered around a shared table. With the emergence and proliferation of digital photography, photo browsing generally happens one-on-one with the computer, and the face-to-face interactions and informal storytelling that used to take place around picture

browsing have been completely lost. Yet what we keep noticing as we distribute our pictures electronically over remote networks and file sharing programs, is that most of our friends and family members would still much rather be shown a picture in person, accompanied by a spontaneously told story, rather than pointed to a URL where they can download the picture by themselves. In short, we have today the power of great digital content, but we're still really stuck at the interface. In order to ensure that we don't lose the appeal of sharing and browsing our growing digital media collections, we need to provide interfaces that can bring back some of the face-to-face sociability and fun into these interactions. After all, there is little reason to accumulate such large media collections if we do not have strong reasons that compel us to go back and look at them.

Tabletop Media Browsing

To address this question of interface in what concerns digital media browsing and sharing, let's take a moment to consider the way we use tables. We sit around them in our homes, classrooms, public spaces and work environments to do many different kinds of things together, such as eat, chat, tell stories, share pictures, or draw. Extending these ideas into the digital realm, we might imagine creating multi-purpose media tables that could tap the wealth of digital content and applications that exist in cyberspace and open up possibilities for people to engage in highly sociable face-to-face interactions around them. On a desktop interface, each single person remains completely in control of their own little portal into the digital realm. In contrast, casual and sociable interactions in the physical world are characterized by shared control between multiple people, where different people can each do different things at once or where they can share in the same thing at the same time. Media tables that provide multiple points-of-control for meaningful simultaneous multi-user interactions within a digital media application context have the potential to enhance the way in which we jointly share and interact with digital media.

Figure 1:

Fred, Amy and Jack browse travel photos and tell stories around the TViews media table.

Imagine a photo browser that runs on a digital media table, allowing friends and family to share their digital photographs in a manner similar to way they used to share physical photographs. Not only can this application bring back the casual face-to-face interactions into digital photo-browsing, but it can tap the power of digital content storage and retrieval to enhance the activity in ways that were not formerly possible. For instance, the use of geographic metadata tags could allow the photos to be automatically organized across the surface of a map depicting the location where they were taken. Perhaps users could then link to additional information resources about the area or connect to remote users who visited the same places and uploaded their images to shared media servers. The following interaction sequence is taken from a photo-browsing session on a media table called TViews that we have built in our lab. This particular dialog involves three people using three interactive objects called pucks to navigate travel photos on a map of Ireland, as shown in Figure 1. In this particular case, Fred and Jack went on a trip to Ireland together and are using the TViews map browser to explore their shared photo collection. They are also

showing their photos to Amy, who was unable to accompany them on their trip and is seeing the photos for the first time.

Fred: [Moving puck to new location]

Amy: [Leaning over to see Fred's pictures but still holding her puck] *Oh, the new paint!*

Jack: [Focusing on his own puck at the other side of the table] *Mmm hmm…*

Amy: [Taking her puck and moving it aside to look at Fred's picture] *That's a nice color.*

Jack: *You hadn't seen that yet?*

Amy: *No.*

Fred: *It looks just like the outside.*

Amy: [Focusing back to her puck at a different location] *Matt...*

Jack: [Looking at the picture Amy is indicating] *Matt Malloy.*

Amy: *Malloy's.*

Fred: [Switching focus from his own puck to Amy's puck] *Mom was very happy to see Mr. Malloy.*

Amy: [Switching locations] *We've already been to that location...* [Switching locations again] *Are you in this one? Yes, I see…*

Jack: [Looking at the new location of Amy's puck] *I like some of these ones right here.*

Fred: [Looking over] *Ah yes, driving on the sand bar, haha…*

Amy: [Focusing back to Fred's pictures for a moment] *I'm sure you're proud of your work.*

Jack: [Pointing at picture from the puck Amy is holding] *Do you see this one right here?*

Amy: [Looking back to her puck to see what Jack is showing] *Yes.*

Jack: *This was an island off the coast and we were driving by over here* [indicating location on map] *and we saw some cars driving over and we were wondering what was going on over there. And mom said "oh, they do auto commercials on the beach."* [Amy takes another puck and drags aside the image Jack is talking about for a closer look] *So we drove closer and it turned out that when the tide is out you can actually drive… you can't really see it in this picture because it's zoomed out…* [Pointing at picture again] *but there are road signs that go all the way through.*

Amy: [Leaning over to look] *Oh weird!*

Jack: *And it's basically what the locals do… they drive out onto the beach but you can only do it at certain times of day.* [Stepping through pictures on puck] *The other picture might actually show it better here…* [Stopping at picture] *That one, see… there's a road sign.*

Amy: Oh yeah, neat!

Even in this short sequence, we clearly see how people are able to fluidly switch focus between their own puck interactions and those of the others around them. The nature of their interactions transitions easily from the way they used to browse physical photographs on a table, and allows casual conversation and informal storytelling to emerge around digital media content. This example shows how a traditional activity like photo-browsing can transition easily onto a novel interaction platform. It also demonstrates that a novel interaction platform like a media table can transition easily into existing spaces like the home living room by making use of and adapting to existing social practices and conventions within the shared environment.

In the remainder of this paper, we take a closer look at the role played by interfaces in our digital media interactions. In particular, we examine how a digital media table can be used to manage and access large-scale personal media collections in shared physical contexts. The discussion focuses on a version of the TViews table designed for the home environment, and on two different applications that were developed for media content management and browsing in this setting. We also discuss the results of a qualitative user evaluation of TViews within a real-world home, which illustrates its usability in the consumer space and has helped us identify potential mappings between physical interaction objects and large media content archives.

The Limits of Existing Digital Interfaces

Today, the notion that our interactions with the digital world are necessarily mediated through different kinds of interfaces is part of mainstream knowledge. These interfaces often make use of physical-world metaphors that translate raw data into representations that we can easily understand and manipulate. The desktop metaphor currently used by most personal computers as a means of organizing interactions and information is an obvious example. It draws from the interactions we have with documents and filing systems at our physical desks to create familiar digital world analogues. Yet while digital information itself is inherently malleable, our interactions with this information are constrained by the interfaces presented to us by specific interactive environments and applications. The controls and displays of digital information shape the way we understand it and what we are able to do with it. As a result, the way in which we understand the computer as an appliance or tool or even as a social object is directly tied to the current state of digital interface technologies rather than to their full or future potential.

Human-computer interaction researcher Bill Buxton has long argued that the current design of computing systems is hampered not so much by issues of speed and power, but rather by interfaces that are poorly matched to different users' skills and contexts and by the weakness of a "one size fits all" approach to computer interface design [Buxton 1996]. Both work and leisure activities in our society are increasingly converging around desktop and laptop computers and employ the same tools and metaphors of graphical user interfaces. Yet the needs of different users vary greatly and necessarily depend on the particular tasks they wish to accomplish and the locations where they are typically performed.

To address the limitations of existing computer interfaces, the past decade of human-computer interaction research has shown a growing interest in emerging areas such as ubiquitous and pervasive computing, and tangible user interfaces [Weiser 1991, Ishii 1997]. Rather than moving more and more of our daily tasks out of the physical world into the limited interaction space provided by desktop computer interfaces, these domains seek to seamlessly integrate the digital and physical worlds in order to enable sensory-rich interactions with digital information within a broad range of contexts and environments. Tangible interfaces in particular, are characterized by the coupling of controls and representations of digital information within manipulable physical artifacts, surfaces and spaces [Ullmer 2002]. This interaction technique differs significantly from mainstream computer interfaces such as keyboards and mice, which act as input devices alone and are intended as controls for the visual representations of digital information provided by screen-based displays.

In order to develop a tangible computing platform for a social space such as a typical home living room, there are a number of considerations that designers must take into account. Clearly, the living room is a very different environment from the office or study spaces for which today's desktop PCs are designed. While the latter two spaces are physically suited to single person productivity-oriented tasks at a desk, the living room is a social setting where far more spontaneous and face-to-face interaction takes place. Sitting at a desk, people interact primarily with information, be it in digital or paper-based form. For this reason, computer interfaces that sit on people's desks focus on enabling seamless interactions between single users and digital information. This information can still be shared with other users, but the sharing happens remotely via networks of personal computers and servers. In social spaces such as a living room, the focus of information exchange shifts from person-to-computer to person-to-person. Some media interactions that take place in a living room can be quite passive, such as watching television or listening to

the radio. But even these often invite conversations amongst the people in the room. At other times, the living room becomes a lively space as people interact both with media and with each other while playing games or exchanging stories. A computer platform designed for the living space needs to take all of this into consideration by accommodating shared interactions with media that do not restrict the control and display of information to a single person alone. The following section describes the TViews media table, a multi-user tabletop computing platform designed for shared spaces.

TViews Media Table and Applications

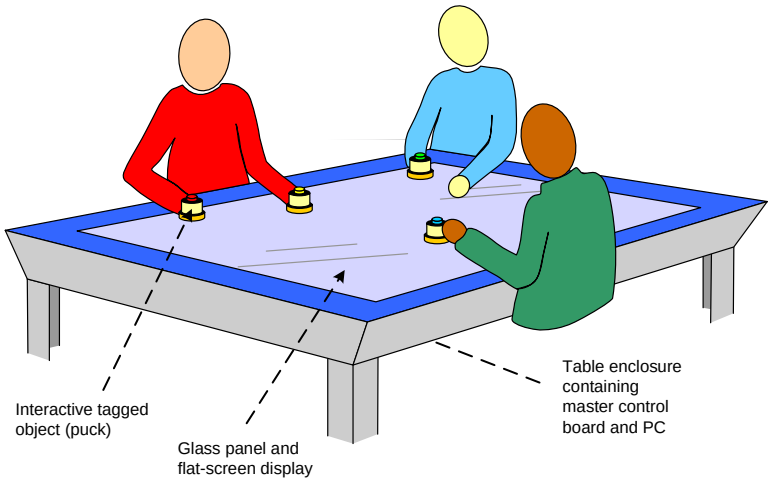

Figure 2:

The TViews media table is based on custom acoustic sensing technology that works through a glass panel mounted above the table's embedded display surface.

A media table is a horizontal surface upon which the spatial configuration of tagged objects is computationally interpreted and then augmented with coincident visual output. The visuals are usually provided by rear or front projection. At our research lab, we have developed a media table called TViews, based on custom-built acoustic sensing technology that works through the surface of a glass panel placed above an LCD screen embedded in the surface of the table, as shown in Figure 2. This technology allows the table to scale to different sizes, and enables an extensible set of interactive objects called pucks to be tracked on its surface and moved from one table platform to another. The approach allows multiple people to interact together in both physical and distributed space, and enables them to have natural interactions with digital media in their leisure time.

Figure 3:

Two working prototypes of the TViews table have been constructed at the MIT Media Lab (left) and at Samsung Research (right).

TViews provides a scalable platform framework that defines a way in which many different media table applications can run on the same platform or on many connected platforms at once. TViews interactive objects can be associated

to a particular application based on their unique functionality or customized physical shape, or used as generic controls for multiple applications. The functionality of the pucks can be extended via externally attached input/output devices, such as buttons, lights or small displays. The current generic puck design includes a top-mounted button and snap-in acrylic pieces that enable color-customization.

The compact sensing and display design of the TViews table allows it to be setup in everyday living environments where there is little or no support for infrastructure that is external to the table itself. To date, two fully working versions of the TViews table have been constructed for living room spaces, as depicted in Figure 3. We describe here two applications that have been developed for media content management and geographic media browsing on the TViews table, shown in Figure 4.

Picture Sorter

Since the first commercially manufactured photo cameras became available in the mid 19th century, the photographic process has achieved mainstream use across the world as a means of everyday image capture. Photographs record the passing moments and memorable events in people's lives, allowing them to live on as visual traces that are permanently solidified in the physical material of the photographic paper. Gathered in photo albums or collections, people can return to them time and time again. This personal repository of memories can be organized, re-organized, browsed and shared together with friends and family members many times. Often, the activity of creating the family photo album itself becomes a joint activity that involves sorting, selecting and grouping the images from a larger collection, and in the process elicits countless reminiscences and spontaneous acts of storytelling.

Over the past decade, we have moved into the era of digital photography. Rolls of film are quickly being replaced with ever-increasing capacities of digital storage, and the lengthy process of analog processing that generates viewable photographs from rolls of film negatives has been replaced with the instant gratification of immediately viewable images. These photos can be shared via mobile messaging, downloaded to laptop and desktop PCs, uploaded to picture sharing websites online, or simply printed out in regular paper form. This transformation in the photographic process allows people to quickly capture far greater numbers of images, but it also requires better tools for organizing these large image collections. Many such tools exist, such as Apple's iPhoto or Google's Picasa, but managing large image collections can be a tedious and repetitive task as it involves manually entering keywords on a computer with no face-to-face interactions between people. The leisurely group activity of sorting through a pile of physical photographs on a table has been replaced with single-user point-and-click interfaces for image organization.

The TViews picture sorting application demonstrates how a tabletop platform can bring back some of the shared fun into the process of organizing digital photographs. The application makes use of the metaphor of a shoebox of images that is dumped onto a table surface for sorting. The images show up in a pile at the center of the TViews table, and the pucks are used to sort them into smaller clusters. Unlike physical photographs which can only ever be in one pile at a time, the digital nature of the application allows the same photograph to be sorted into several different clusters. While the application currently provides only basic sorting functionality, users will eventually be able to add annotations for their photographs via text-entry interaction objects, allowing the system to make use of intelligent sorting strategies based on commonsense reasoning techniques. Future versions will also allow image clusters to be saved as slideshows, posted to the web for remote access, or viewed within different browsing views such as inside the map browser discussed in the following section.

Map Browser

As digital photography continues to spread and evolve, camera technologies are becoming more and more intelligent, and the information they capture along with the visual image itself allows software tools to automatically classify and retrieve images based on their metadata tags.

Nowadays, most digital cameras store the date and time as well as a variety of additional information, such as the aperture, shutter speed and other camera settings that were used when the image was taken. It is also possible to attach GPS (Global Positioning System) receivers to some digital cameras in order to record geographic location information, and in January 2005 Ricoh released the first digital camera with a fully integrated GPS receiver, the Ricoh Pro G3. Certain cameras, including many of the camera-equipped cell phones, allow users to add voice annotations to their photographs. One can imagine that in the not-so-distant future digital cameras will also incorporate sensors to keep track of ambient or environmental information such as weather conditions.

The TViews map browser organizes images on a spatial map based on the time and location at which they were taken. A timeline on one side of the display surface is color-coded by date. Small colored dots at different locations on the map indicate the dates on which photographs were taken there. Users can attach the pucks to different dates by dragging them over the timeline. Once a date is attached to a puck, it can be dragged around the map to reveal images at different geographic locations. The images appear clustered around the puck, and can be zoomed by pressing the puck's top-mounted button. When an image is zoomed, another puck can be used to grab hold of it and drag it to another part of the table for separate viewing.

Figure 4:

Picture sorting and geographic picture browsing applications running on the TViews table.

TViews in the Home

A preliminary user trial was conducted to examine the TViews table in use within the home environment. The goal of the study was to see how people engaged in shared interactions around the table using the media content management and browsing applications described above, and to gather their feedback and suggestions towards future development. The informal and situated nature of this preliminary evaluation allowed us to observe how one particular home environment assimilated the physicality and playful intent of the table, drawing the focus of participants away from the novelty of the platform.

We selected a principal volunteer who was willing to offer the use of his living room for the purpose of our testing and evaluation. The volunteer was a young professional who lives in a one bedroom apartment in Cambridge, Massachusetts. The table was moved to the apartment for a several week period, during which small groups of people were invited to interact with the table. In addition to the principal volunteer, sixteen other people came to try out the table typically in groups of two or three at a time. Given the small size of the apartment, larger groups were difficult to accommodate. Most of the participants were digitally savvy young professionals and students aged 25-35. Their professions ranged across technology-related fields (project management, consulting, software engineering) to landscape architecture, law, library services, publishing and marketing.

The interaction sessions were conducted in an informal manner as small social gatherings during which people could try out the media browsing applications on the TViews table. The apartment owner acted as a host for the social events, and the invited guests were for the most part friends or family members that he sees on a regular basis. The

host created a relaxed atmosphere for the interaction sessions, and participants could play with the table in a realistic setting rather than in a contrived laboratory space. Since the participants were mostly young professionals with full-time jobs, the interaction sessions were typically held in the evening or on weekends, and generally involved an informal dinner followed by casual chatting and social interaction in the living room, where participants sat on a couch and chairs around the TViews table, jointly browsing pictures and sharing stories as shown in Figure 5.

Figure 5:

Users browse picture collections in the TViews map browser application.

Results and Discussion

Over the course of the user interaction sessions, we engaged people in informal brainstorming to gather their feedback about the table and applications. We used video and audio recording for data collection purposes, and participants were asked to complete a questionnaire before leaving, in order to gather background information about each participant's level of media and technology use as well as specific comments about the interaction sessions.

The pictures used in the Picture Sorter and Map Browser were from a recent trip that the host took to Ireland with some of his family members. People adjusted quickly to the use of multiple pucks to browse and sort the images, and much of the time two or more people would be viewing different collections at the same time. We observed that when this was happening, people were able to fluidly and repeatedly switch their focus between their own interaction and that of others. This seamless switch in focus is common when people browse physical photographs on a table, and demonstrates that this natural type of interaction transitions easily into the tangible tabletop space. We also observed a lot of shared interactions, where one person would be controlling a puck, and the other person would use a second puck to drag images from the first person's collection in order to take a closer look at them. Again, this shows that the TViews table provides people with the ability to interact in a similar way with digitally stored pictures as they would with physical ones.

The map browser was the more successful of the two applications. People generally don't want to organize photo collections that are not especially relevant to their own lives. In this particular case, they had the option of browsing the images on a geographic map that provided a physical and spatial context for the collection. By framing a photo collection with context (provided in this case by the application) and story (provided in this case by the host), even an unfamiliar photo collection can suddenly become fun to browse. The host told many stories about his travels around Ireland, while his family and friends navigated the pucks around the map, asked questions, and searched for particularly interesting photos. Some people suggested that the sorting application and map browsers should be paired together as two different views of the same collection – one for organizing and editing the collection, and the other for searching, browsing and viewing the images.

In both applications, large numbers of images were sometimes difficult to handle when they were all displayed at once. This was more problematic in the Picture Sorter, since all of the images remained visible at all times. In the

Map Browser, only the particular days or locations that a user was viewing were visible at any given time, which worked well as an organizational strategy. One person suggested that a zooming approach in the Picture Sorter could be used to zoom in and out of different clusters of images in order to support multiple different views or examine only subsets of the image collection at a time. In this way, it would also be possible to support multiple levels of clustering, where sub-clusters could be created within higher-level clusters.

This question of how to visually represent and manage large collections of pictures or other kinds of data on the screen is an issue that comes up in traditional graphical user interface design as well. Potential strategies include hierarchical views and spatial layouts that can be panned or zoomed. For media table interface design, there is an additional parameter that needs to be considered – that of directionality. Currently all the pictures on the TViews table face in one direction only (i.e. the normal orientation of the flat-screen display in the table). Yet people often sit on all sides of the table when they interact, and so text and pictures appear upside down or sideways to certain users and can be difficult to view. The Personal Digital Historian system developed for the multi-touch surface DiamondTouch from MERL presents an interesting solution to the problem, in which information artifacts projected on the table are automatically oriented to face the nearest edge [Shen 2003].

Another question that was raised for the Picture Sorter was how to annotate or tag the image clusters. Since our media table does not currently provide a keyboard device, text input can be problematic. Several ideas came up during the brainstorming sessions. Some people suggested using an on-screen keyboard where the letters could be selected using the pucks. Another idea was to provide a keypad-device as an add-on for certain pucks that could be used in the same way that a cell phone keypad is used for text messaging. One user also suggested that it would be nice to be able to use a normal cell phone to send text messages to the table or to specific pucks on the table.

Overall, participants found the picture management and browsing interactions to be both intuitive and engaging. Users acknowledged the potential for media tables in the home to provide a completely new way of visualizing large media collections on a shared display surface, and they appreciated the way in which this sociable interaction is very different from what they experience using desktop PCs. They were especially excited about the way in which this type of shared media platform could allow the same kind of casual social interactions and informal storytelling to arise not just around a small set of images from a family photo album, but around a repository of media shared across the world. We found these comments to be gratifying since they indicate potential for a real change in the way we browse and share digital media with in our daily living spaces.

Towards the Future

The interfaces that shape our digital interactions change over time, often independently from the underlying algorithmic constructs and data representations that drive them. Across the history of digital computing, the interfaces presented to users have taken on a variety of forms: from the card punch and reader, to the keyboard, mouse and vertical screen of personal desktop systems, to the small screen and keypad of modern-day handheld devices. These and other interfaces have shaped the way we perceive and understand the realm of computing and the types of operations and interactions we can perform within it. While the range of applications covered by these interfaces is broad, the types of interactions they afford between people are limited to those that are necessarily mediated through the technology.

As our personal media collections grow ever-larger, we need to find strategies that will engage us with them over time. We can entice people to revisit their archived memories and past experiences through novel kinds of interaction platforms that can exist in their daily living spaces and become a part of their lives as social beings. Through these kinds of interfaces, media collections can become more personally meaningful, and can weave themselves into our everyday lives across both the digital and physical worlds.

References

[Buxton 1996] Buxton, William (1996) "Absorbing and Squeezing Out: On Sponges and Ubiquitous Computing" in *Proceedings of the International Broadcasting Symposium*, November 13-16, Tokyo, pp.91-96.

[Ishii 1997] Ishii, H., Ullmer, B. (1997) "Tangible Bits: Towards Seamless Interfaces Between People, Bits and Atoms" in *Proceedings of Conference on Human Factors in Computing Systems (CHI '97)*, ACM Press, pp.234-241.

[Shen 2003] Shen, C., Lesh, N., Vernier, F. (2003) "Personal Digital Historian: Story Sharing Around the Table" in *Interactions*, March/April 2003, pp.15-22.

[Ullmer 2002] Ullmer, B. (2002) *Tangible Interfaces for Manipulating Aggregates of Digital Information*.Ph.D. Dissertation, Massachusetts Institute of Technology, 2002.

[Weiser 1991] Weiser, Mark (1991) "The Computer for the Twenty-First Century" in *Scientific American*, Vol. 265, No. 3, pp.94-104.