about the ...

TECHNOLOGY HORIZONS PROGRAM

The Technology Horizons Program combines a deep understanding of technology and societal forces to identify and evaluate discontinuities and innovations in the next three to ten years. We help organizations develop insights and strategic tools to better position themselves for the future. Our approach to technology forecasting is unique—we put humans in the middle of our forecasts. Understanding humans as consumers, workers, householders, and community members allows IFTF to help companies look beyond technical feasibility to identify the value in new technologies, forecast adoption and diffusion patterns, and discover new market opportunities and threats.

INSTITUTE FOR THE FUTURE

The Institute for the Future is an independent, nonprofit strategic research group with more than 35 years of forecasting experience. The core of our work is identifying emerging trends and discontinuities that will transform global society and the global marketplace. We provide our members with insights into business strategy, design process, innovation, and social dilemmas. Our research generates the foresight needed to create insights that lead to action. Our research spans a broad territory of deeply transformative trends, from health and health care to technology, the workplace, and human identity. The Institute for the Future is located in Palo Alto, California.
# Contents

1 Introduction .................................................. 1

2 A Spectrum of Technologies .................................. 3
   Communication: From Closed to Open ...................... 4
   Sensing: From Simple to Smart .............................. 6
   Sensemaking: From Formal to Informal .................... 8

3 A Spectrum of Applications .................................. 13
   Exploring Places: Walled Gardens vs. Open Guides .... 4
   Communities: Smart Homes vs. Aware Grids ............. 6
   Shopping: Smart Stores vs. Activist Shoppers .......... 8
   Public Spaces: Surveillance vs. Sousveillance .......... 10
   Highways: Smart Roads vs. Self-Aware Traffic ......... 12
   Desktops: Software Guides vs. Software Matches ...... 14

4 A Spectrum of Impacts ........................................ 29

Appendix A Spectrum of Viewpoints .......................... 35
   Deborah Estrin
   Instrumenting the World: Sensors, Networks, and Power 36
   Michael Goodchild
   Geospatial Awareness: Finding Meaning in Places ...... 39
   James Landay
   Activity Models: Embedding Computing in Everyday Life 43
   Dan Brickley
   The Decorated World: Ontologies Meet Folksonomies .... 48
   Clay Shirky
   Groups and Tags: The Social Context of Context Awareness 52
In 1988, Mark Weiser laid the foundation for what he called the third wave of computing.

The first wave was mainframe computing, followed by the second wave of desktop computing. The third wave, according to Weiser and others, would be a kind of ubiquitous computing—or calm computing, as Weiser called it—in which technology would recede into the background of our lives.

The world is now on the brink of this third wave, and at the heart of it is something we might call context awareness. In the simplest terms, context awareness just means having information about the immediate situation—the people, roles, activities, times, places, devices, and software that define the situation. But context awareness is also about meaning and meaning-making, and it is especially this piece of the technological puzzle—this sensemaking—that we would like to address in this report, The Many Faces of Context Awareness: A Spectrum of Technologies, Applications, and Impacts (SR-1014).

Over the last year, we interviewed many of the leading innovators in context-aware technologies, and we uncovered as many definitions of context awareness as we have experts to speak about it. The visions of context-aware applications range from 24-hour-a-day records of our life experiences to appliances that recognize, monitor, and advise us. And real-world applications are already beginning to appear in the form of everything from simple air-bag actuators in our cars to buddy lists that signal our presence and tagging systems that provide a way for groups to describe shared space, whether it’s physical or virtual.
To understand this range of technologies, viewpoints, and applications, we have developed a spectrum of context awareness, ranging from those that emphasize top-down design on one end to bottom-up emergence on the other. We hope that this framework will help you anticipate the kinds of successes and failures we’ll see in this new wave of computing, as well as the social implications as computing moves into the environment.
The context awareness of the next decade will evolve out of interaction of three key technologies: wireless communications, sensors, and semantic technologies (or sensemaking technologies). Within each of these technologies, there is a spectrum that will begin to give shape to range of context-aware applications we can expect. Yet the dimensions of the spectrum for each technology differ.

- Wireless communication technologies will be defined primarily by their openness.
- Sensors will be defined by their smartness.
- Semantic technologies will be defined by their degree of formality.

The following pages array the important technologies of context awareness across these three spectrums.
At the closed end of the communications spectrum are traditional phone, mobile, and cellular services. The service model here is to sell fixed bandwidth network connections, routing, and switching services. Over time, these network operators may serve as “trunk carriers”—almost as backbone services for ad hoc open networks at the other end of the spectrum.

As we move toward the middle of the spectrum, we find the newer standards for wireless services like Wi-Fi and Bluetooth. There are few commercial operators of Wi-Fi and Bluetooth services. Access points and connections may be deployed by anyone flexibly, one connection at a time. The spectrum here is unregulated.

ZigBee is a set of high-level communication protocols designed to use small, low-power digital radios as a network for sensors. Like Wi-Fi and Bluetooth, Zigbee networks can be built very cheaply and deployed one node at a time. Large networks, with large numbers of devices and a large coverage area can form autonomously and operate reliably for years without any operator intervention.

ZigBee, Wi-Fi, and Bluetooth can all be deployed in combination with fixed closed network services, as well as completely open user-configured networks.
Software-defined radios (SDRs) are radios that provide software control of a variety of modulation techniques, wideband or narrowband operation, communications security functions (such as frequency hopping), and waveform requirements of current and evolving standards over a broad frequency range. SDR-enabled user devices and network equipment can be dynamically programmed in software to reconfigure their characteristics for better performance, richer feature sets, and advanced new services that provide choices for each end user.

Ultra-wideband (UWB) radios can change frequencies contextually by transmitting very short duration pulses, often only nanoseconds or less, across a wide spectrum of radio frequencies. UWB is sometimes described as spread-spectrum or frequency hopping.

Mesh routing enables each device to route messages autonomously to other devices, bypassing closed networks completely. Mesh networking is a way to route data, voice, and instructions between nodes, allowing continuous connections and reconfiguration around blocked paths by “hopping” from node to node until a connection can be established—becoming what David Reed describes as “networks without carriers.”

Mesh routing creates “networks without carriers.”

Source: Wireless Ad-hoc and Local Area Networks Research Lab, North Carolina State University.
Sensors are devices that can detect and respond to light, movement, moisture, temperature, chemicals, electronic charges, or radio signals. In their simplest forms, sensors trigger some kind of switch or simply record a response.

Even the simplest of sensors can, however, play an important role in building context-aware systems. Today’s passive radio frequency identification (RFID) tags, are inert, with no power of their own. They simply reflect a fixed, stored digital number only when in the presence of an external radio source—or RFID reader. Yet they, or their technological offspring, will play a growing role in instrumenting the physical world.

Location sensing is a bit more complicated type of sensing. Global positioning system (GPS) chips actively listen to the timing signal of several distant satellites and then calculate latitude, longitude, and altitude. Similarly Wi-Fi receivers with proper software can listen for unique base-station addresses, compare them with a locally cached database of known base stations, and then calculate and report location, just like GPS.

Bluetooth devices may also begin to function as sensors, as applications read Bluetooth signals and even automatically read data streams from them for specific purposes.
In the middle of the spectrum, some hybrid devices are emerging. For example, RFID devices are beginning to be repurposed to serve stored data, such as location coordinates.

Perhaps more dramatic are WISPs from Intel’s Wireless Identification and Sensing Project at its Seattle lab. WISPs combine RFID with particular kinds of sensors, such as accelerometers, to achieve both sensing and communications capability in a very inexpensive device that requires no battery. Like all RFID chips, WISPs are activated by a radio signal, but instead of returning fixed data, a WISP chip can send specific contextual data using RFID data formats.

Very smart sensing devices are able to sense environmental conditions and then make simple analyses and communicate, either on request or autonomously and dynamically. Motes™, from Dust, Inc., are already being deployed in the field. Each mote can include a sensor, a computer, memory, and a radio, all on very small device, a few centimeters square. Deborah Estrin at the University of California, Los Angeles describes a future of intelligent independently powered labs-on-chip that can gather and process information about specific environmental context at any given location and then forward the information to other nodes in the network, autonomously.

Similarly, microscale and eventually nanoscale robotic devices will be able sense phenomena, gather and process data and take action based on the micro-local context. In time, Estrin forecasts devices that will allow machine learning to occur across multi-sensor systems.

**WISPs combine sensors with RFID to provide a cheaper, no-battery alternative to motes.**

*Source: Intel Research Seattle*
The tools at the formal end of the sensemaking spectrum include standard vocabularies for describing context, as well as taxonomies (formal classification and categorization systems), and ontologies (maps of the relationships among objects or terms in a formal vocabulary).

For example, medical information is described using standard medical terms from a system called the Unified Medical Language System (UMLS), developed by the National Library of Medicine. The UMLS meta-thesaurus is a map for machines and people to translate and understand medical context from across dozens of medical disciplines and subdomains.

The open-standard geodata vocabularies GML/WFS are another example of a formal vocabulary—in this case, for linking all kinds of data to particular places.

Activity models are a somewhat more emergent form of sensemaking, describing the context of activities. As currently under development by James Landay at the Intel Research Seattle lab, activity modeling uses Web crawlers to find procedural information on the Web—that is how-to instructions. These are automatically converted into a set of relationships among physical objects. These relationships can then be mapped on to sensor data about real-world objects in a particular setting to determine what activities are underway in that setting.
The Semantic Web is a formal ontological approach to describing documents. Proponents of the Semantic Web hope these document descriptions will become the standard format for data interchange in the future, replacing today’s file formats. Many groups have created standard vocabularies for describing documents in their domains.

At the same time, attempts are underway to bridge these formal vocabularies with informal tagging, potentially making them more open and emergent. Dan Brickley, at the W3C, points to the practice of faceting as a way “to articulate a road map where these [ontological] structures can bubble up from scruffy, messy, grassroots data sets.” Faceting is a practice from library science that analyzes repeating patterns in a vocabulary in order to define them as explicit relationships.

Also, massive amounts of data from user-defined domains like databases are coming online and could be harvested for contextual applications. The XML family of knowledge structures developed by W3C has the built-in capability for such data objects to self-identify—and thus serve as contextual data.

At the most informal end of the spectrum are the emergent practices of applications like Flickr and del.icio.us, where users assign their own tags to personal or shared data. When it occurs, coordination is achieved primarily by sharing lists of existing tags, which serve as folksonomies—or informal, emergent taxonomies.

At the same time, aggregates of these tags, viewed as tag clouds, can be analyzed for valuable meta-information, such as the order of tagging, the amount of daily activity associated with particular tags, and the daily content domains of individual groups. These views of group activity provide a new window on previously hidden behavior and will undoubtedly lead to new kinds of social and economic activity.
Semantic processing represents a host of new—and not so new—technologies for inferring meaning from data. With roots in artificial intelligence, the formal approaches to semantic processing tend to develop comprehensive, top–down views of knowledge domains. These ontologies are often displayed as tree-like structures and define, a priori, a set of data categories as well as the relationships among those categories.

Perhaps one of the most ambitious of these ontologies is the CYC Project led by Doug Lenat in Cycorp. CYC is an attempt to create a model of all human knowledge. After 20 years of continued hand coding of “common sense” human language, large-scale, robust CYC systems are still in development at Cycorp and apparently in use for contextual applications, primarily for military, intelligence, and government clientele. Smaller beta versions of CYC are released from time to time on the Web for experiments and research. These can be downloaded for free at http://www.opencyc.org.

At the other end of the spectrum, in an attempt to make sense of the vast stores of metadata emerging from bottom–up tagging of blogs, photos, Web sites, and even physical places, researchers are developing techniques for automatically creating, analyzing, and applying tag clouds.

A tag cloud is essentially a bottom–up map of the relationships among tags (which, in turn, represent the contents of some data collection). The relationships are often represented by order of creation, by size of the collection the tag points to, and/or by the links among tags. For example, the image on this page is a tag cloud that presents a snapshot of IFTF’s collective set of del.icio.us tags on a day shortly before our Fall 2005 conference on context awareness.

**CYC represents all of human knowledge, at a very high level, with this ontological map.**

**A del.icio.us tag cloud—a snapshot of IFTF’s thoughts on context-aware technologies.**
Another example of a tag cloud is the graph net created by Ciro Cattuto; using simple code, Cattuto is able to build a network visualization of the inferred semantic relationships among the tags that Flickr users attach to their photos.

Note that both of systems are dynamic: they are automatically updated moment by moment, using the changing base of tags that they scan. They also foreshadow a future where highly intelligent context-aware software will be able to search, translate, combine, and correlate data from both formal and informal descriptions of context by searching and comparing graph models of linked knowledge.

**A graph net representation of the tags attached to Flickr photos.**

**SOME KEY DEFINITIONS**

**Semantic:** relating to the study of meaning; semantic processing is the analysis of data to determine its meaning.

**Metadata:** data about data; a library card catalog contains all the metadata for the library’s collection.

**Ontology:** a map of the relationships among pieces in a collection; a family tree is an ontological description of kinship.

**Taxonomy:** a formal classification system for complex sets of objects; a botanical taxonomy might classify plants according to their genus and species.

**Folksonomy:** an informal classification system emerging from free-form tagging of data and objects; a list of all the Flickr tags you share with friends or family is a folksonomy of your photographs.

**Tag cloud:** an informal map of the relationships among metadata tags in a data collection; tag clouds are typically represented as visual network diagrams, using the size of the nodes to indicate the frequency of use of a particular tag.
The spectrum of context-aware technologies will provide a new set of tools to develop new applications that help users better navigate tasks and places—and change the dynamics of our social interactions. Like the technologies that support them, these applications will range along a spectrum from top–down, design-directed tools and activities to bottom–up, emergent phenomena.

In this chapter, we examine this range of possibilities with a set of six scenarios, each with a top–down and bottom–up implementation. Each scenario focuses on a particular domain of activity or experience, including:

- Travel
- Home and community life
- Shopping
- Public spaces and public behavior
- Commuting
- The desktop
One of the oldest visions of context-aware computing has been the location-aware travel guide. Combining orientation, navigation, and contextual information retrieval, these tools will combine the sensing functionality of a *Star Trek* tricorder with the knowledge and clarity of a great guidebook.

However, the opposite ends of the context-aware spectrum are providing two very different means of delivering contextually useful information to travelers.

From the top down, wireless communications companies are delivering pre-packaged content about points of interest throughout their service areas. Coupled with handsets that limit the access of third-party developers to the user’s location, mobile carriers are pursuing so-called “walled-garden” strategies as they strive to recoup the massive cost of broadband buildout. The goal of these strategies is to own and therefore license the user’s contextual experiences of places via mobile devices.

*Verizon’s Mobile “Get-It Now” service provides location-based information from Verizon mobile phones.*
Meanwhile, open network standards like Bluetooth and Wi-Fi are providing an alternate means of delivering relevant information to users without tapping into the service provider’s own location sensing data. In September 2005, Aalborg, Denmark became the first local government worldwide to provide information and updates to visitors’ mobile phones using Bluetooth beacons positioned throughout the city’s streets and sidewalks.

The MIT Media Lab’s History Unwired project also uses Bluetooth to re-script the established tourist geography of cities by luring PDA-equipped visitors to less well-known historical sites that are truly “off the beaten path.”

Meanwhile, bottom–up wiki tools allow users to create their own online guides to places. For example, the Vegan Guide to Oxford is a wiki with user-created reviews of restaurants in Oxford. Eventually, this kind of user-created guide is likely to be combined with Bluetooth location sensing to provide truly bottom–up guides.

The “Vegan Guide to Oxford” is a wiki with user-created reviews of restaurants in Oxford.

Source: The Vegan Guide to Oxford...
Residential neighborhoods will be an important venue for context-aware computing.

Today, most companies think of the smart home as a place that does things for us to make our lives better. In this very top-down view, the context-aware home will track people and objects throughout the house and proactively respond to changing situations and needs. The home may turn lights and appliances on and off, alert us to dangers, or just alert us when the washing machine has finished its cycle. The PlaceLab project at MIT’s School of Architecture, for instance, is experimenting with application scenarios in an off-campus apartment that has been extensively retrofitted with sensor grids.

One compelling application for the smart home is elder care. Smart homes are being targeted as a way to allow seniors to age in place, sensing what activities they are engaged in, assisting them in those activities, and letting remote caretakers know how they’re getting along throughout the day. Intel Research expects to be able to develop such a system, using WISP technology and activity models, for under $5,000.

*Intel’s CareNet display provides remote, wireless information about the well-being of seniors living on their own.*

Source: Intel Research
However, another version of the smart home is developing from a broader definition of the home as part of a regional infrastructure complex. The fragile, over-burdened electrical power grid is an ideal candidate for this enhancement.

Researchers at Pacific Northwest National Laboratory are developing smart appliances that can sense the health of the power grid by measuring the frequency of AC power coming into the home and shut down temporarily during periods of peak use. Future versions will be able to request power before it is needed, allowing the grid to forecast demand and schedule delivery of power. Instead of focusing inwardly on what is going on inside the home, aware grids draw upon sensors in the home to develop bottom-up understanding of infrastructure use at a very fine grain.
For most of us, stores will be the first place we encounter context-aware environments in our everyday lives.

In the new few years, context-aware technologies will rapidly proliferate in retail shopping environments. Smart stores will use various tags and visual cues to identify individual shoppers and provide relevant information about products. One powerful vision of this scenario was portrayed in the science fiction film *Minority Report* (2002)—wall-mounted retinal scanners triggered personalized holographic video advertisements directed at passing shoppers.

Outside science fiction, early indicators can be found in the smart store experiments of Europe and Asia. Metro AG, Europe’s largest retailer, operates its “Future Store” in Rheinberg, Germany. The store combines a variety of context-aware applications aimed at consumers:

- An RFID-equipped Smart Card that lets shoppers download their shopping list directly to a cart-mounted display
- Smart shelves that allow the store to provide you directions and availability to items on your list
- Automated self check-out

*Metro AG’s Future Store in Rheinberg, Germany, uses sensors to streamline shopping and integrate user needs with promotional opportunities.*
These top–down context-aware systems are prompting not only consumer protests over privacy concerns, but also new bottom–up context-aware applications. Ironically, context-aware technologies will allow shoppers to arm themselves with powerful sensing and sensemaking tools of their own. They will be able to retrieve independent reviews of products, comparative pricing from other vendors, and product safety and nutritional information.

One example of this kind of bottom–up context-awareness in stores of the future comes from the TIVIK project in Finland. The service allows shoppers to snap a photo of a product’s bar code using a mobile phone and send it over a 3G wireless network, fetching information about price, nutrition, quality, and safety. For food items, the system even estimates how much exercise is needed to burn the calories in a serving, based on the user’s height, weight, and age. Users will also be able to develop shopping and food consumption logs that can later be mined for personal health and fitness analysis.

With the TIVIK Project, shoppers can read barcodes and get wireless information about price, nutrition, quality, and safety of products as they shop.

Source: http://virtual.vtt.fi/tivik/
New sensing and sensemaking technologies—from simple RFID tags to video face recognition—are being widely used to improve surveillance of public spaces. Firms like Identix are marketing highly capable systems that can match faces captured by surveillance cameras against photographic databases in real-time. Today, these tools are largely employed by governments to monitor airports, sporting venues, and other critical infrastructure. In the future, such top-down, aware surveillance systems will be deployed by other large institutions like corporations, hospitals, and universities to micro-manage access to campuses, buildings, and rooms.

Identix is one of several firms that offer real-time face recognition service to agencies that monitor public spaces.
Sousveillance, a term coined by Steve Mann, is the inverse of surveillance. It refers to the role of the individual—often an activist—in recording and documenting the abuses of those in control or in power. For instance, numerous instances of citizen-reporters providing accounts from New Orleans after Hurricane Katrina conflicted with and questioned official reports.

However, this notion of sousveillance can go far beyond political activism. According to Marc Davis, Director of Yahoo! Research’s Berkeley lab, future forms of sousveillance will increasingly be used as a form of peer-to-peer public behavior control. For example, in June 2005, a woman on the subway in Seoul, South Korea neglected to clean up after her dog defecated on the floor of one of the train cars. A disgusted onlooker snapped a high-resolution photo and posted it to a popular Korean photo-sharing site. The image quickly became the focus of national outrage; the woman was identified, shamed, and even expelled from her university.

*This wearable wireless webcam is a sousveillance device designed to mimic surveillance cameras used in public spaces.*

Source: http://wearcam.org/
While human users are at the center of many emerging forms of context-aware computing, future application scenarios are also likely to include machines that exchange contextual information, independently, without human intervention.

Some of the oldest venues for context-aware computing can be found in factories, where automated assembly lines employ robots that sense and react to a constantly changing material context. Such networks embody the principles of top–down context-awareness: they employ passive sensors and communicate using rigid vocabularies over closed networks.

*Today, most assembly lines, like this chip fabrication facility, depend on varying degrees of automation.*
In the future, we can imagine another extreme—where independent robotic devices will negotiate contextual exchanges with other familiar and perhaps unfamiliar robots in what might be described as a “robot ecology.”

Current research aimed at creating “swarm intelligence” among groups of autonomous robots, such as the GridSwarm project at the University of Essex, are developing the basic techniques that will enable this kind of production in the future. The GridSwarm project uses Linux and tiny Bluetooth connected computer modules, embedded in robotic airplanes and helicopters, to perform parallel distributed computing tasks while flying in flocking formations. Ultimately such capability could be used to by small robotic devices to self-organize to perform complex tasks in the physical world.
Contrasting approaches to context-awareness will dominate our passage along roads and highways of the future.

Transportation planners and highway authorities have long deployed surveillance cameras and embedded roadway sensors to monitor traffic flows. However, these grids are now being expanded and linked to traffic information systems to provide contextual information to influence driver behavior. For instance, California’s Performance Monitoring System (PeMS) collects data from thousands of sensors embedded in the San Francisco Bay Area’s freeway network. It then develops estimates of travel time, which are relayed to drivers approaching key crossings and interchanges via large roadside displays.

*California’s PeMS project uses sensor data to display traffic congestion.*

Source: Performance Monitoring System
In contrast to sensory road infrastructure, car manufacturers are experimenting with bottom–up approaches to context-awareness on the highways by augmenting vehicles with wireless communications capabilities.

Cars already contain many sensing instruments for things like vehicle speed, exhaust, and ambient temperature. By adding the ability for cars to exchange information with other nearby vehicles, and relay this information across distances from one vehicle to the next, researchers envision a scenario in which warnings of slowdowns ahead can propagate rapidly back through traffic, providing a similar service to top–down systems like PeMS.

Still, car companies are not the only advocates of bottom–up intelligence on the road. The “Roadcasting” project at Carnegie Mellon University has developed a platform for sharing music wirelessly between cars stuck in traffic. The project is described as collaborative mobile radio. The system can read preferences of those around them and choose songs or podcasts that those people want to hear on their own car stereos or wireless devices.

The Roadcasting Project prototypes an application in which cars share music wirelessly.

Source: Roadcasting Project, Carnegie Mellon University
One of the earliest we saw places context-awareness implemented was in desktop computing. However, context-aware principles remain at the margins of today’s user interfaces.

For instance, there are plenty of software “guides” that observe user’s actions and then suggest productivity hints based on that context. Microsoft’s “Clippy” is perhaps the most notorious and reviled example of the poor state of this kind of context-awareness in desktop computing. In the future, rather, we will see a shift toward context-aware software that is designed to minimize user distraction and assist “attention conservation.”

*Microsoft’s Clippy was a top–down effort to define user context and offer advice—whether it was wanted or not!*

Source: Microsoft
In a contrast to the top-down, predetermined context-aware software guides are applications that create context awareness through searching and pattern matching.

For example, the Social Network and Relationship Finder (SNARF) is a prototype application developed by Microsoft researchers to automatically perform triage on incoming e-mail. It analyzes new and old e-mail content to infer the relative importance of different correspondents.

Another example is Google’s AdSense system, which analyzes search terms and matches them with indexed Web content and advertisers who bid for the right to display relevant advertisements or have their Web sites come up at the top of the page.

These point to more emergent forms of context-aware software.
The spectrum of technologies and applications presented in this report suggest a range of scenarios for the future development of context awareness, with a range of social and economic impacts.

More important, however, is the dynamic between the two ends of the spectrum and the very different paradigms they each represent, both technologically and socially. It is perhaps this dynamic—sometimes in harmony and sometimes in conflict—that will have the largest impact on both innovation and social discord over the coming two decades. The following chapter outlines how we see this spectrum of these impacts.
At the top–down end of the spectrum, information about context is designed to be predictable and homogeneous, and it comes from specified sources. Here, context awareness will be all about easing the burden of daily life, and people will look to technology to make life more productive, more precise, more secure, and more convenient. Their goals and actions will coalesce around the following trends:

- **Delegating responsibility to agents.** Humans will trust software agents and machines to organize and execute more of their tasks.

- **Replacing human functions.** Technology will be tasked with providing better services than humans can provide at lower cost.

- **Maintaining secure spaces.** It will be possible to more completely secure spaces that are not easily enclosed by walls, fences, and other boundary markers.

- **Instrumenting processes.** More routine and repetitive processes, including very complex processes, will be measured, monitored, and automated using contextual data.

- **Alerting.** People will rely on context-aware tools to prompt them when they need to act, effectively outsourcing some of their traditional scanning and filtering functions.

- **Predicting.** Analysis of contextual data will increase human ability to forecast complex and long-term phenomena.

In short, innovations at this end of the spectrum will provide more control with less effort.
At the bottom–up side, context awareness emerges from human social behavior. It is more unpredictable. It comes from lots of sources, and it may be socially filtered. Here the goals are often aesthetic and emotional, and the solutions are profoundly human-centric. Specifically, the tools and practices at this end of the spectrum will emphasize:

- **Expanding human awareness.** As more and more information is connected to places, objects, and social interactions, human awareness of our complex reality will grow.

- **Extending human capacity.** Rather than replacing human functioning, context-aware technologies will extend our cognitive and social capabilities, enhancing memory and mental functioning as well as connecting us to much larger networks of people.

- **Decorating the world.** With context-aware tools, the world becomes a palette for self expression, signaling a general shift from consumption to creation.

- **Building group identity.** The tools of context awareness will support group discovery, group production, group-defined contexts, and low-cost group coordination.

- **Valuing innovation and disruption.** Riding waves of new awareness, new meaning, and new social connections, people will “surf” the physical world in the coming decade the way they surfed the Internet in 1990s, extolling both innovation and disruption as virtues.

In short, context awareness at the emergent, bottom–up end of the spectrum keeps humans alive and active in creating meaning in a data-enriched world.
The two ends of our spectrum are likely to co-evolve, each feeding the other with new opportunities for innovation. As Dan Brickley, a member of the Technical Staff at the Worldwide Web Consortium (W3C), suggested, informal tags and folksonomies will be harvested automatically for more formal structures of meaning—and more top–down applications of datamining. And more formal platforms of context awareness will equally give rise to new opportunities to “hack” these systems for social and expressive innovations at the bottom–up side of the spectrum.

This creative co-evolution suggests a scenario of rapid innovation and expansion of context awareness, with both predictable and unexpected opportunities for creating economic value. However, the same dynamic—where top–down control meets bottom–up hacking—may also create darker scenarios when the interests of those at opposite ends of the spectrum come into conflict.
This clash is perhaps most evident in the evolution of surveillance and security in the face of
smart-mob behavior. As cities embed security surveillance systems on their streets, people who
don’t want to be to be monitored will develop bottom–up tools to “sense the sensors” and move
about the city unseen. Top–down applications for crowd control, public safety, and even smart
consumer displays will encounter activist shoppers and contested aware cities—places where con-
text-aware technology is at the center of a cat-and-mouse game of detection and counter detection,
action and avoidance. For example, a group of artists in New York City calls itself the Institute for
Applied Autonomy and has mapped the location of surveillance cameras throughout Manhattan.
The group has created a Web site where anyone can plot “a path of least surveillance” between
two points.

But the clash will also play out in the marketplace as open-source models, peer production, and
cannibalization of existing networks draw economic battle lines between top–down strategists
and bottom–up hackers. In the most disruptive scenarios, this clash will play out in the increasing
criminalization of emergent innovations that threaten the top–down patterns, with a corresponding
growth of black market and “dark mob” activity as the emergent end of the spectrum goes under-
ground.

The challenge of the next decade, then, will be to avoid the polarization of the world of context awareness
and seek a rich ecology of approaches with new strategies—and perhaps institutions—for mediating the
objectives of each.
As part of the research process for this project, we consulted more than a dozen people with expertise in context-aware technologies. This included academics, corporate researchers, and industry professionals, and we gathered their input through an expert workshop and one-on-one interviews. The thoughts of the experts contributed greatly to the depth and richness of the forecasts in this report. Here we present the highlights from a few of the interviews we conducted to give you some deeper insight into the ideas the experts shared with us.

The Experts

< Chris Beckman, Researcher, Intel Research Berkeley
< Dan Brickley, Member of the Technical Staff, Worldwide Web Consortium (W3C)
< Marc Davis, Founding Director Yahoo! Research Berkeley
< Deborah Estrin, Director of the Center for Embedded Network Sensing (CENS), University of California, Los Angeles
< Scott Fisher, Chair, Interactive Media Division, School of Cinema-Television, University of Southern California
< Markus Fromherz, Manager, Intelligent Systems Lab, PARC
< Michael Goodchild, Director of the Center for Spatially Integrated Social Science, University of California, Santa Barbara and Associate Director, Alexandria Digital Library Project
< Quentin Jones, Associate Professor, New Jersey Institute of Technology
< James Landay, Director of Intel Research, Seattle Lab
< David Pescovitz, Co-Editor BoingBoing.net and Research Affiliate, Institute for the Future
< Greg Pottie, Professor, Electrical Engineering, University of California, Los Angeles
< Ted Selker, Professor, Context-Aware Lab, MIT Media Lab
< Clay Shirky, Professor, Interactive Telecommunications Program, New York University
< Terry Winograd, Professor, Computer Science, Stanford University
DEBORAH ESTRIN

Instrumenting the World: Sensors, Networks, and Power

Deborah Estrin is Director of the Center for Embedded Network Sensing (CENS) at University of California, Los Angeles. Focusing on large-scale, distributed systems composed of smart sensors and actuators that are embedded in the physical world, CENS is developing the underlying theory and practical experiments for building a world in which intelligence is distributed not just through a virtual world, like the Internet, but through the physical world as well.

In this interview, Deborah begins by taking a long view of the emerging world of nanotechnology and how it will extend the capacity of microscale devices. The microscale devices themselves will move out of the lab and into the field as autonomous instruments, but instruments that can be integrated into systems via network technologies. While the communication and basic sensing technologies for these autonomous systems are already well-developed, the challenge exists in what Deborah calls machine-learning and adaptive systems—systems that can alter the data they are collecting and focus their sensing capabilities based on feedback from the environment they’re sensing.

According to Deborah, “the real effectiveness and intelligence and adaptability of these systems is going to come from advanced statistical techniques for dealing with this very disparate, lousy, and uncertainty-plagued data. … Everything in these systems is going to be multi-scale, and you’re going to have little, local, simpler filters that are being reprogrammed by higher-level, more sophisticated filters elsewhere in the system.”

On Context-Aware Observation Systems

On the IT side of things, I see sensing as having an important impact in a wide range of science, engineering, consumer, civilian, and government activities as we build our ability to observe and control physical worlds.

In the ten-year timeframe, I think we’ll have pervasive, distributed observation systems, and we’ll be starting to move toward systems that also are somewhat involved in actuation and in mediation. So if you start out thinking about being able to monitor slopes for possible failure as a result of heavy rains, if you talk about monitoring water reservoirs to identify the much finer spatial and temporal water quality for health risks, I would certainly expect to see that happening in these next ten years. I think as we get to the end of that timeframe and move into the 20-year timeframe, you’ll start to have automated remediation.

On Context-Aware Technology Moving from the Lab to the Field

There’s a big story about DNA on a chip, but right now, that’s still going on in the analytical context of a laboratory. It’s not literally that you put a chip out in a water reservoir, and it autonomously does sample preparation and analysis and identification. We’re not there; they’re just starting to look at those kinds of things in the lab. So the lab-on-a-chip happens in a lab. It really makes the analytical context much more productive and efficient, but it’s still not something that is an autonomous instrument.
In the next ten years, though, the autonomous instruments will move—thanks to MEMS and, in some cases, to nano-scale devices—from the laboratory where you have analytical instruments doing sample prep and things like that to really being able to do that in the field. And that’s where we’re focusing our work right now related mostly to MEMS. Sensory development today is trying to make those sensors fully autonomous by creating integrated systems. And those things take a long time to gestate from original idea to actually getting the engineering right and the process of building them. So that’s a good five or ten years before that stuff hits the field.

If somebody comes up with a business model for more active RFIDs—not just the passive funky reader and very, very short-range, passive tags—but if, in the active tag arena, a business model emerges and drives a significant investment, the cost and capability would be greatly improved, and I think that could be a significant disruption.

Precision agriculture and things related to agriculture are the most certain uses of this technology. There’s no question that that is happening and will happen. But how large these uses are and whether they’re large enough to push a market, I don’t know.

On Human Assistants, Not Autonomous Agents

Let me just mention a mistake that I initially made in looking at this technology. I was overly focused on completely autonomous systems. But in fact, a tremendous amount of the power of these systems in the coming decade (and I would think even longer term) is going to come from seeing them as human assistants.

The analogy I like to use is this: what if physicians had decided that CAT scan or MRI technology was only really usable when the images could be automatically interpreted? Of course, that’s not the case; these tools are used as observing mechanisms that human beings then interpret, and they have truly revolutionized medicine.

So the same goes for the technology of context-awareness and embedded sensing. It can have tremendous power in providing 3D surround-sound observability, even though the final interpretation that might be done in a very interactive way by a human being.

On Semantics: Machine Learning and Sensor Modalities

Another key set of developments is related to machine learning and techniques that let us really deal with uncertainty and with combining multiple sensor modalities, multi-scale information sets, lousy information, and very diverse sorts of information.

Improvements here could radically change the effectiveness of our systems. No matter how densely we think we’re sampling the world relative to traditional times or other technology, we’re always under-sampling. And the real effectiveness and intelligence and adaptability of these systems is going to come from advanced statistical techniques for dealing with this very disparate, lousy, and uncertainty-plagued data. Multi-scale, multi-modal sensor information processing is such an important component of this kind of context awareness, and there’s so much intellectual work to be done there.
Eventually, these systems are going to be adaptive in terms of what data they’re collecting and where they’re focusing their sensing capabilities. There’s going to be a huge amount of internal adaptive behavior in these systems, driven by machine learning, even when what you’re doing is providing a very powerful observing system to a human being. It’s like having adaptive objects inside your telescope. Or again, going back to the medical analogy, as MRI and CAT scan and imaging technologies have improved, part of what’s improved then is the internal adaptation, which improves their ability to resolve images and things like that.

**On Articulation, Actuation, and Adaptive Detection**

Articulation and actuation are going to be important functions of these adaptive systems. If you think about what you can see through a camera, if you can just move it, if it has some degree of freedom, it’s a qualitative difference in terms of the numbers of images you need of a particular scene or thing or the number sensors you need to get multiple perspectives on it.

So these systems will be able to do this kind of internal autonomous articulation, and to get you your best information, this articulation is also going to have to be driven by what people think of now as statistical computing.

**On the Future Trajectory of Sensing Technology**

In the next 5 to 10 years, it’s not just chemical and biological DNA-type analysis that will transform our ability to observe. It’s also the ability to embed automatic processing of images and acoustic input. We can call a microphone a sensor and we can call an imager a sensor, but for the most part, those are really very raw data-gatherers that require a significant amount of human interpretation to actually make an observation.

I think we will see embeddable imagers as sensors, programmed with pattern recognition and event detection and even adaptive event detection so that these things are really acting as sensors in the optical domain—and similarly in the acoustic context.

**On Sensors with Onboard Adaptive Filtering and Pattern Recognition**

You’re going to have simpler things right on board, and then those things will have to be adjusted in the context of more context, if you will. Somewhat higher-end devices are seeing things that will allow these sensors to adjust the local filters that are running. Everything in these systems is going to be multi-scale, and you’re going to have little, local, simpler filters that are being reprogrammed by higher-level, more sophisticated filters elsewhere in the system. The challenge here is not just about computation, it’s about statistics—meaning that it’s the statistical technique, not just the ability to crunch.
Michael Goodchild is the Director of the Center for Spatially Integrated Social Science at the University of California, Santa Barbara. He is also Associate Director of the Alexandria Digital Library Project, which has been working for more than a decade to create a distributed digital library of material that is “geo-referenced.”

Michael points out the problems of multiple collections of geodata, multiple standards for coding geodata, the lack of any coding at all on much of the data that currently describes place, and an overall decline in the collection of geodata. He sees promise, though, in commercial frameworks for geodata like Google Earth, Keyhole, and Microsoft’s Virtual Earth, as well in recent innovations in amateur and bottom-up geocoding, such as Fundrace.org, which geocoded political contributions, and Metacarta.com, which is building the kind of knowledge base needed to distinguish Moscow, Idaho, from Moscow, Russia.

For Michael, the key challenge at present is that “there’s been very little work on getting to the point where you can send a request, not for all the information relevant to this location in Santa Barbara, but rather specifically for the elevation of this location. … And that ought to be something that is totally doable.”

One of the challenges of implementing a context-aware geospatial web is the many different ways in which geodata has been coded and stored in collections.

So how do you know which portal to choose? We’ve done some work on CLM [collection-level metadata] content standards, and that’s part of the answer.

There’s also the Federal Geospatial Data Committee, which has led to ISO standards for metadata. The issue there, though, is that those are object-level metadata. They describe an individual object, and so they’re fine for supporting the search mechanisms within a site, but they don’t resolve the question of how to choose the site in the first place.

But unless there’s a significant global effort and unless it’s possible to persuade every lower level in the administrative hierarchy, like the State of California, to use a global one-stop, then you will never get a single point of entry; you will never get a single portal. And as long as we know that to be true, there’s got to be some way of choosing among the alternatives.

I think we’re still in this phase of believing that it’s possible to be all things to all people. So we’re still building the California geospatial data warehouse on the assumption that anyone needing geospatial data in California will automatically go to that portal.

And that assumption simply isn’t true. It’s not based on a comprehensive definition of geospatial, so typically what appears in the warehouse is only a small subset of all geospatial data. And it doesn’t address the needs of someone in Santa Barbara who is looking across county-, state-, or federal-level sources, and it doesn’t address the question of division of the world by theme, either.
As a result, when I need to find geospatial data, I need to make a complex set of judgments about the area covered by the data I’m looking for, by the type of data, by the theme of data, and based on that, I make a guess as to which warehouse is most likely to contain the data. But I think the assumption behind the construction of many of these things is still that it’s possible to be all things to all people.

**On Building a Geospatial Search Engine**

In principle, one could build an engine that would find all geospatial data and automatically harvest it. We fall short of that because a lot of spatial data is still not self-descriptive and a lot of the formats don’t include the necessary metadata tags. For example, ESRI shape files don’t have to have metadata associated with them. But in principle, they could. And so just like Google harvests essentially all text, one could harvest all geospatial data.

I think that Keyhole and Microsoft’s Virtual Earth, and Google Earth are actually quite significant. They provide a uniform framework that removes a lot of the issues that the GIS community has struggled with: there are no variations in format or variations within the structure in their models. Everything is resolved to a single, unified grid. It’s a hierarchical grid, so you’ve got options with respect to resolution, but at least you’ve got a single structure.

**On Deep Versus Specific Information About Place**

None of the key technologies is getting to a level of responding to a query. There’s been very little work on getting to the point where you can send a request, not for all the information relevant to this location in Santa Barbara, but rather specifically for the elevation of this location.

To answer than kind of question, you’ve got to identify a data set and then you’ve got to enter the data set and make a query against the data set. Think about library metaphor. Instead of saying that the library’s responsibility is to put a book in your hands, the library actually needs to read the book for you and answer the questions you want to ask.

And that ought to be something that is totally doable. But I think almost all of the effort to date has really been on the first aspect of giving you the data set, from which you then can, potentially, answer the question.

**On Geospatial Metadata: The Semantic Web**

If we could get everybody to agree to use Geographic Markup Language (GML), that would be one solution to creating interoperable collections of geodata. Using a small finite number of data formats, one could build engines that would harvest the metadata from all of those. The problem is, of course, that we haven’t made that agreement. The GML strategy doesn’t really deal with the dominance of certain commercial players. It’s a data exchange format, but a data sets won’t necessarily exist in an exchangeable format.

The problem is how to get people to create metadata for their own data. A lot of geographers and cartographers don’t even label their data at all. So getting them to provide object-level metadata is a struggle. A lot of effort has gone into trying to persuade people to create metadata. But here’s typically not that much in it for the custodian of the data.
On the Limits of Sharing

One of the assumptions we have is that data are worth sharing, and I think that assumption needs to be opened up. One could ask, “Is there any value in sharing at all—and if so, over what domain?”

Often, it’s solely the domain of the creator of the data; in effect, he’s the only person interested in sharing in any way. Or it may be a much broader domain, but unless that domain is defined, it’s very difficult to make any objective decisions about how to describe the data—what kind of metadata fields to use.

I’m not at all optimistic about the future of automated sharing of data. I think it’s not at all clear why somebody in Santa Barbara, California, would want to share data with somebody in Reno, Nevada. And the circumstances under which that might occur are very limited.

On the Problem of the Data Supply

I think one could make a very strong argument that it’s actually the data supply that we need to be worried about, rather than the data sharing question. The construction of data warehouses hasn’t really changed the nature of the data supply, and in many ways, the data supply is actually deteriorating.

We are comparatively worse off now than we’ve been in previous decades. In fact, the peak of mapping of the planet probably occurred in the 1950s and 1960s. So I think there are a lot of other questions that maybe ought to take priority in terms of making it possible for people to do context-aware computing. And one of them is what kind of data they’re going to be able to rely on.

On Amateur Geocoding and Locative Media

There’s been some stunning progress here recently. Fundrace.org is one of my favorite examples: somebody simply took the database of political contributions in November and geocoded it, and then put a map interface on the front so that you can just create a map of your neighborhood.

This kind of thing is becoming endemic, and it’s really taken off. I think another kind of activity is Metacarta.com. It’s really done very well in terms of building the knowledge base that’s needed to disambiguate Moscow, Idaho from Moscow, Russia, or figure out whether Shanghai is a verb or a noun.

On Mobile Interfaces

The assumption that the industry’s been making is that the mobile interface is very similar, but more constrained. And I think that’s maybe a little bit too narrow.

I’m not sure that the average mobile user really needs or can use a map. A map is a rich source of information, but it’s not necessarily the best source when you’re working through a very constrained user interface. So in the vehicle, for example, spoken driving directions are probably much more useful to the average user than a map display. And similar things may occur with cell phone use.

So I think we’ve got some way to go in thinking through the question of user interfaces for mobile devices. In general, the whole field has been constrained by a lack of good theory. I think we’re still in the business of retrofitting a theory behind a technology that’s got ahead of the theory. And that’s
really in sharp contrast to a lot of other fields where the theory existed before the technology developed, such as mathematics or statistics.

With stronger theoretical concepts, we’d be able to simplify the field enormously. I think we spend a lot of time worrying about things that, in principle, we shouldn’t. Map projections are an example. There’s an enormous complexity in the field because of the need to flatten the earth, and yet the earth essentially doesn’t need to be flattened when it’s seen in a digital environment. That’s a legacy of paper. So I think we’re still struggling with a paper legacy and with a field that really doesn’t have as strong a theoretical framework as it should.
JAMES LANDAY

Activity Models: Embedding Computing in Everyday Life

James Landay is Director of Intel Research at its Seattle lab, where his team is working to define an underlying strategy that takes advantage of embedded sensors to provide support in daily life. In particular, James is interested in the potential of context-aware systems to support the elderly as they “age in place” with tools that make it safer and easier for them to live alone, but integrated into a care network that has up-to-date information about how they’re doing.

The underlying theory for James and his team is activity theory, and their main strategy is the use of activity models that help interpret what a person is doing from the way they interact with the objects in their environment. The team uses automatic web crawlers to find how-to information for tens of thousands of activities and then creates ad hoc computer models of those activities that can be compared with data reports from sensors in the environment. The lab has also innovated the sensor technology, developing the WISP sensors described in Chapter 2 on “A Spectrum of Technologies.”

James emphasizes that it’s all automatic: “the models don’t have to be great. They just have to be good enough. Once you have the models with real people doing the real activities, you can actually refine the models from real data, again, automatically, without the person having to be involved ... that’s how we got to the point where we have 40,000 models today.”

Definition of Context Awareness

We think about a future of computing that’s ubiquitous, in the Mark Weiser sense: embedded in our everyday lives and supporting them. This kind of computing has to understand more about our context if it’s not going to be overly intrusive. If we have to actually manipulate this computing the way we need to manipulate our desktop computers and our laptops today, then we actually won’t use it for much more than we use them today. So if we want to get to a point where we’re going to have a lot more computing supporting our everyday lives, it has to understand context.

And then the question is: what is context? To me, context is about those bookreport questions that you answer in junior high: who, what, where, when, and maybe how.


Some of those questions are really easy, I think. For the “Who?” question, we have lots of solutions to tell us who someone is. If you’re carrying a computational device like a cell phone or a PDA, we typically know who you are. It’s your device. It’s a personal device. Of course, it gets a little harder when you’re in a public environment, and the computation is embedded somewhere else. In that case, we might have to do something smarter to figure out who you are. “Where?” is another question people in our lab have worked on—figuring out the location of someone using the computational devices that they might be carrying.

The interesting questions that are perhaps going to be most important are “What?” and maybe
“How?” So what is someone doing in the physical world? What’s the activity? By understanding their activities, we can better support people in those activities. Without that understanding, I think computation won’t move much beyond what it does today on the desktop or simple communications devices. It will remain a fairly low-level type of support; it won’t have much offer at a high level.

So a lot of the research going on in my lab is about supporting activities—both simple physical activities, such as “I’m cooking right now,” and also the high-level activity that you might think of if you think of activity theory in the Russian psychology sense.

In this sense, activity has a high-level objective, such as, “Let Grandma live a healthy and active life.” It also has a set of subjects: the people who are helping Grandma live that active and healthy life and the tools that are mediating between those people and that objective.

So this is what I see as the big missing element in what someone would call context-aware or even ubiquitous computing today. It hasn’t been solved because it has a lot of the hard artificial intelligence problems associated with it. But I think when this problem is solved—and I think we’re actually on the cusp of solving a chunk of it here—it’s going to create a big change in the way we think of user interfaces and computing.

On Hierarchies of Activity: Modeling Context

In activity theory, activities function in a hierarchy. At the lowest level is something called an operation, which is kind of an unconscious thing that someone is doing. For example, when you know how to type and you’re just typing away, that’s really an operation. Or when you know how to drive a stick-shift car, shifting the clutch and the shift stick is just unconscious, and that’s an operation.

The next level up really is what we call an action, and that’s similar to what people in human–computer interaction generally refer to as a task. So an action might arise if you’re trying to pass someone on the highway and you might actually think: “I’ve got to go around that guy on the left,” and then you’re going to shift properly and accelerate, which is unconscious.

And then a high-level activity is something even more removed. It might be: “I’m trying to get from my home to my mom’s house.” Or even, “I’m just trying to stay close to my family.”

Most computing technology easily supports low-level operations. We know how to automate those tasks. But when you start to get to the activity level, it starts to get harder. It’s hard for computing devices to know that you’re cooking.

On Supporting High-Level Activity

The next level is trying to understand the higher level activity: oh, you’re taking care of your health or you’re eating or you’re exercising. That’s the AI research we’re doing in our lab now, and we’ve made a lot of progress in it. The new research that I’m doing is trying to learn how to use that activity paradigm as a new way to think about user interfaces and applications in general. Think of building applications to support high-level activity, not to support tasks.
Now think of some of the successful new computing devices out there today. One reason they’re successful is because they support high-level activity. For example, TiVo supports the idea that I want to watch what I like right now versus the VCR model, in which the user is focused on recording this channel at this time. And there’s this gap between “record this channel at this time” and the activity that the person actually wants to engage in, which is: I want to watch something that I want to watch right now. The gap between the two actually makes the VCR a hard-to-use device, and the reason TiVo is not hard to use is that there is very little gap there.

**On Emergent Model Building**

One of the problems with any of these kinds of machine-learning approaches to inferring what people are doing is building the models. The previous solution was to hire an AI post-doc or grad student and let them crunch out three to five models and give them a Ph.D. But we need to make thousands of things, at least at this middle-action level, so that’s approach is too slow.

What we do instead is we build models automatically by just crawling the Web. So far we have built 40,000 models this way. Originally, we did this is by finding certain Web sites that have step-by-step instructions on how to accomplish some task or activity. For example, cooking sites have recipes that are essentially step-by-step instructions on how to cook a certain item, and others sites are more general “how-to” sites. For the tea model it says, “Boil water.” And then it says, “Add a teabag. Add some sweetening and some milk.” And by screen-scraping that recipe, I actually can figure out what the objects are that you need to interact with and in what order. And I don’t care about the verbs so much. All I have to know is, oh, there’s a teapot, there’s water, there’s sugar, there’s milk, and maybe the ordering—the temporal constraints between them. And I suddenly have a model.

More recently, we’ve been able to build these models without relying on those how-to sites. We crawl the top 100 sites for a certain activity and capture the nouns automatically.

Then the technical way to tell which objects someone is touching is to put RFID tags on objects, which we think are going to become ubiquitous on lots of objects soon. If not ubiquitous, they’ll be so cheap that you can attach them for situations where you want this high-level support. And then we have an RFID reader that’s in a bracelet, so as you touch objects we know what you’re touching. And we can also do this with another kind of special RFID tag we’ve built called a WISP. That’s an RFID tag that has a sensor in it and communicates that it’s moving. It’s totally battery-less. It gets all its power from the RFID radio signal. And so with a long-range RFID reader, you can tell which objects I’m touching in what order and infer what I’m doing.

**On Context-Aware Elder Care Technology:**

When you have elders today who are getting frail or have early-stage Alzheimer’s, people try to let them live in their homes as long as possible. But at some point, they can’t, and they have to go into some kind of care facility. That’s really a negative scenario for several reasons. First, the quality of life is really compromised when they have to go live in this facility and can no longer be around their own objects, their homes, their friends and family. Another negative impact is simply the cost.
It’s really expensive to put people in these facilities. So we as a society are not going to be able to afford to do this in the long run.

So one of the major projects here is answering the question: how do you support elders aging in place? And the research on this has led to a new product division in Intel on digital health. But one of the applications we could think of from this activity-based perspective is to allow Grandma to live a healthy, active, and independent life. And the starting place is to know which activities Grandma is doing in her home—the activities of daily living. For example, we might want to know that Grandma made and ate breakfast, lunch, and dinner; she took her medications; she got some exercise; and she socialized with her friends.

If we were able to automatically detect those activities we could display that information to their care network—say, Grandma’s eldest daughter, Grandma’s cousin, and her doctor. Then those people could have a better sense of how Grandma is doing and take care of her better without being nosy and having to call her up all the time and treat her like a kid, saying: “Did you take your pills? Did you eat?” Instead, we’ve found in our research that they treat her better and talk about more substantive issues when they talk. And both the elders and the kids like that better.

Now suppose Grandma has early-stage Alzheimer’s and starts to forget what she’s doing while cooking. We can imagine a prompting system in the kitchen that reminds her of the next steps. It can also remind her to take her medication. And again, that same system is built on an understanding of the activities that Grandma is engaged in—the high-level activities as well as the tasks and actions that she’s taken.

We can also infer Grandma’s physical activity because we’ve built devices that Grandma can put into a cell phone or a small clip-on device that can tell whether Grandma is walking, running, bicycling, taking the stairs, going inside/outside, sitting on the couch, standing, and so on. Putting all this data together with activity inferencing, along with a user-interface design that support these activities across a set of users who are playing different roles—that’s the kind of application I think is the future of computing and is not well-supported by today’s paradigms or technology. Yet this activity perspective makes it possible.

You will probably be able to build this elder-care system, to have someone come into the home and set it up, for $5,000 in five years or less.

On Context-Aware Technology’s Design Challenge: Multiple Devices, Applications, and Users

I think one of the big problems is that, to do it right, you’re actually building a system of things that are working with different devices and what we think of today as multiple applications. And today, most people don’t design for that. You design this Web site or you design Microsoft Word. You design this new laptop. What we’re doing is almost more like systems integration. And today, the only people who routinely design this way are defense contractors. Most designers for the commercial space don’t do it at all. So we are going to need tools as well as education to help people design for this kind of future.

In fact, a lot of the manufacturers try to put up barriers to this kind of broad-based integration because they want you to buy their stuff. But for embedded computing to be successful, all the pieces
will have to work together. We’re already seeing this, though, with things like Universal Plug and Play and the Digital Living Network Alliance. They’re making standards for home-computing devices and applications to work together. Those kinds of efforts are going to have to accelerate.

From a technical perspective, processes will need to adapt automatically to work across different devices because you won’t know exactly which device your client has. If I’m making the software for Grandma for this elder care, yeah, I want it to run on this special portrait display. But if they have a Web browser or they have a phone, I want it to work there too. And I don’t want to have to develop 20 versions of this to work across all those things.

**On Adaptive User Interfaces**

Also, as people use these systems over the long term, they will find new uses; they will become skilled and adapt themselves. So our systems need to be able to improve and change over time. For example, as I learn to drive that manual transmission car, shifting becomes an automatic operation rather than a high-level action that I have to think about. And it’s similar with these interfaces. Over time people are going to become skilled and fluent.

So I think that’s going to be one of the big breakthroughs—a paradigm that allows fluent use to occur. And I think this whole activity-theory perspective supports that. It’s a move away from the cognitive science that has dominated how we think of human–computer interaction, toward something like activity theory, which is based on early 20th-century Russian psychology. It doesn’t have all the answers, but it’s a different perspective. Activity inferencing is the key that’s been missing.

Now that we can tell which activities somebody is engaged in, we can apply some of these theories to make computation both more valuable and easier to use at the same time.
Dan Brickley has spent several years working on the Technical Staff of the Worldwide Web Consortium (W3C), specifically to develop the concepts and standards for a semantic Web. More recently, he has created the Friend of a Friend (FOAF) Project to create a Web of machine-readable homepages that describe “people, the links between, and the things they do and create.”

In our interview with Dan, he described the context-aware world as a “decorated world. He began by envisioning a future in which today’s file formats are replaced by documents that conform to the Resource Description Framework (RDF). He went on to reflect on the state-of-the-art of metadata and, in particular, on the tension between formal ontologies and informal folksonomies.

He concludes that “there’s no real opposition between the folksonomic approach and the ontological approach, but we certainly need to articulate a road map where these [ontological] structures can bubble up from scruffy, messy grassroots data sets. And then as you find yourself suddenly with 20 million records on your hands, how can you go back and say, all the stuff that was tagged “Bristol” is actually referring to the city in the United Kingdom rather than the city north of Boston?”

Definition of Context Awareness

In my work, the sense of “context-aware” relates to the unavoidable overlaps between tasks, between areas of life or domains or whatever you like to call them. The Semantic Web approach looks at the ways computing applications currently fail us. Our applications are currently very file-format-centric: one format per task. Even with the rise of open document formats based on XML, this heritage has shaped our thinking.

But the world around us just isn’t like that. Everything overlaps awfully. Describing photos, describing geography, describing people, describing events and calendars, describing small businesses, or describing media files—describing all these as separable tasks is something that holds computing back from offering an integrated experience.

How does this relate to context? The idea is that, keyed to some combination of place, task at hand, or object of interest, machines should be able to offer us other relevant avenues to explore. It’s the dream of a decorated world.

On How File Formats Limit Our Ability to Create Context-Aware Applications

If you think about computing applications, they’re organized with one file format per task. You run a piece of software, and it will store its information in the file particular to either that particular piece of software or that class of software. So, spreadsheets might use the dominant spreadsheet format, calendars might use an open standard for calendaring, and so forth.

That’s not a good model for information sharing because the world is full of overlaps. Whenever you think you’re addressing, for example a mapping task, you find yourself tangled up with issues of photography or of describing people or events. The world out there doesn’t come cleanly parti-
tioned, and information sources don’t come nicely tagged as, “Here’s a database about people.” Or, “Here’s a database about events.” Or, “Here’s a database about photos.” You have a database, and each entry in the database will be about, in some sense, people, photos, places, or small businesses.

So what we’ve been working on in the Semantic Web community is a framework for information that mixes across domains at a very fine-grained level. If the Web and network computing are going to make the most of information someone has created in one set of categories, we need a framework for mixing it with stuff that other people have filed under different categories.

By 2015, I think the file format as the foundation of data interchange in the computing world will have gone away, and we’ll have moved to a computing environment where machines exchange documents that use terms defined in Semantic Web dictionaries or Semantic Web ontologies.


There are really two kinds of tensions here: tensions among formal ontologies and tensions between formal and informal descriptions—that is, between ontologies and so-called folksonomies.

So let’s start with the formal ontologies. The Web is the most successful decentralized system, and the architecture of the Web reflects that. It was decentralized in terms of identifiers, in terms of protocols and formats. And that heritage carries through into the Semantic Web work. The approach taken in the Semantic Web design was not to have a single formal ontology defined by a standards committee. We made a technology that allows different groups to play in the same marketplace and to explore different levels of compatibility with their colleagues and with their rivals.

Look at the history of the Dublin Core, for example. The Dublin Core defines a set of properties that are appropriate to online bibliographic retrieval: title, description, date, and so forth. Now, when a certain community of publishers and rights holders came along, they looked at the Dublin Core and published a rather scathing critique, because the Dublin Core has an ontology that comes from a resources-discovery perspective. It simplifies certain things that an ontology for rights holders simply can’t blur. If you have a video, and it has a soundtrack, and the soundtrack has certain rights associated with it, you want to know that; whenever the video is played on television, someone who contributed in a certain named roll to one of its constituent parts needs to be paid. An ontology that captures that richness is a fantastic and a beautiful thing, but it’s not necessarily a priority for people who are creating an ontology for resource discovery.

So even before we get to the distinction between folksonomies and ontologies, we have these series of tensions—and collaborations—between two groups who are both using RDF. They’re both using XML. They’re both using Web technology. But because they have different priorities and different business drivers, they’re creating different ontologies. And one of the challenges that we’ve had in the Semantic Web world is assuring at least a partial understanding between these data formats. So how could a rights-management system that used an ontology index or the work from MPEG7, partially share knowledge with the Dublin Core-based system that was in the library tradition of card catalog records and a thesaurus?
On Folksonomies and Faceting: Emergent Structures of Meaning

There’s a word in the library world—faceting—that applies to thesaurus systems. If you have managed a thesaurus system for a number of years, you end up with a huge tree of concepts related by broader and narrower branches. And what you find, inevitably, is that the principles for partitioning those concepts have regularities that were not initially evident and not necessarily structures that you could have anticipated. So the library community, over the last many years, has come up with conventions for documenting these repeating patterns. For example, you might have a tree of concepts and a regular branching structure based on locality, so the geographic facets of that information can be pulled out and made explicit.

This is something that’s not really found its way into the folksonomy and Web-tagging community yet, because it’s too new. It’s only been going on two, three, or four years as the number of blogs has increased and the number of user-friendly tools like del.icio.us and Flickr have popped up. But if you look at the way people use those sites now, you find tags that contain a lot of hidden structure.

Recently, I’ve been looking at geographic tagging on Flickr, and people have been creating tags that hide a lot of assumptions that could be made more explicit to machines. You might have a tag that says not just “cat,” but a picture of a cat to send to my friend Tom. You might have a tag that says “geocode on the lat equals minus something.” They’re doing all sorts of things, and they’re doing them in a way that makes perfect sense to them and their immediate applications, but it’s very hard for them to communicate to potential consumers of this data that are scattered across the planet.

So, I think there’s no real opposition between the folksonomic approach and the ontological approach, but we certainly need to articulate a road map where these structures can bubble up from scruffy, messy, grassroots data sets. And then as you find yourself suddenly with 20 million records on your hands, how can you go back and say, all the stuff that was tagged “Bristol” is actually referring to the city in the United Kingdom rather than the city north of Boston?

There’s the framework for articulating to machines the possible relationships between scruffy grassroots data and more formal ontologically modeled and hence more expensive data. That’s an engineering task.

But then there’s the data-mining piece—the “best guess” problem perhaps. If we go back to Flickr and look under a particular tag, we’re very likely to find a lot of playful ambiguity, where people have been having great fun overstepping the line or reinterpreting or using metaphors. But we’ll also find mistakes, and the mistakes are there because there’s no prose text in the user’s natural language that explains what that category is supposed to be all about. Mining that stuff is a business opportunity, but it’s also a huge headache.

Data mining something the size of Flickr or something the size of the blog community can go a long way toward showing the implicit structures that govern our communication. But a lot of times people say things that are just purely ambiguous, and in the absence of a machine having common sense, there’s going to be a lot of risk involved in computer guessing.
On the Dream of a Decorated World

You’re in a shop, and you want to see where your fair trade coffee comes from. You want to talk to its producers. You want to see if your T-shirt came from a sweatshop. Or you’re in a bar for lunch, a nice place you’ve just discovered, and you want your calendar to automatically grab the schedule of concerts there.

So many “smart environment” scenarios boil down to the idea of the machine having, at its fingertips, a thousand avenues to explore. The Semantic Web part of this puzzle is making sure there’s an integrated, consistent data model to support the kind of data queries and information filtering that will be needed to flesh out that vision.

In a decade, I think we’ll have a Web of descriptions that apply to pretty much everything of interest, and we’ll have a good handle on the technology for tuning into the descriptions we want, regardless of the domain they were created in, regardless of who created them, appropriate to the task at hand is. The dream of Web annotation applied to the world is what the Semantic Web is about.
Groups and Tags: The Social Context of Context Awareness

Clay Shirky teaches in the Interactive Telecommunications Program at New York University and maintains a list of essays on Network, Economics, and Culture. He has come to be regarded as one of the leading thinkers on the emerging role of groups in a world where group-forming technologies make it easier for groups to discover themselves and work together to achieve coordinated aims, sometimes more effectively than traditional organizations and institutions.

We asked Clay about his well-known essay “Ontologies Is Overated,” and he shared his vision of world in which any set of unique things—URLs, latitude/longitude markers, cell phones, and other device addresses—could all be tagged informally to add meaning to it. He sees a world in which people do this tagging as a by-product of other activities that they’re already doing, and he sees them doing it in groups.

As he thinks about the impacts of this kind of behavior, Clay concludes that “the overall value that I’ve seen across multiple systems is the replacement of planning with coordination … context awareness means that groups can dispose, to some degree with planning in favor of coordination.”

Definition of Context Awareness

Context awareness is a way to de-atomize the view of any given situation. Instead of cutting it up into its atomic elements, you’re starting to be able to say: “Okay, I can understand something about the context in which this material was originally embedded.”

It’s a fundamental change in our ability to extract value from systems rather than items. But the value in a social software application like Dodgeball and the value for a long-haul trucking company are different. Both are going to be relying on the ability put a time and location stamp on a physical object in real-time. But in one case, the physical object is a container filled with gloves destined to an industrial plant. And in the other case, the object is the object of desire—the woman you’re interested in. And the difference in value between those two systems that rely on the same technology is, I think, going to be quite, quite profound.

On Tagging and Tag Clouds

We see a lot happening around tagging right now—which I would describe as moving away from formal categorization toward free-form labeling and using all the advantages of scale essentially to re-create the value of categorization, but at a much lower cost.

One of the interesting things about tagging is that a tag makes almost no sense out of context. Some tags are produced by the people writing the material, those tagging the blog posts and so forth. Other tags are produced by readers. Then there are tags produced to characterize things versus organizing information, people, and objects into groups; or just to signal issues of state (such as “To Read” or “To File”).

And one of the things that’s becoming clear is that when you start comparing tags as used on TiddlyWiki, Flickr, Technorati, and del.icio.us, you have to have context. And the more context
you have, the more valuable the tag becomes. So you want to know not just the tag but the service it came from, the user who used it, other tags that were used at the same time, and so forth. And everything then becomes an interpretive problem of making sense of all of those markers. The context itself is this cloud of metadata that we leave behind, sometimes because we mean to and sometimes not. The new challenge becomes the reading and extracting of value from that cloud.

For example, it turns out that the order in which people enter tags is actually significant. So if you reorder them into alphabetical order, you’re actually destroying value. It’s not just the tags that are important. The classic atomic strategy would be: this is what the tags are. But in addition to the atomic aspect, there’s actually the significance to order. So we have this growing body of tag cloud parsing tools for people trying to figure out how to extract the value that’s obviously present but doesn’t respond to traditional modes of extraction, for things like classic categorization systems.

The Paralog Curve: A Portrait of Group Agreement

The really interesting one–two move with tagging is, first, that anything that’s unique is a logical site for a label. A URL is unique, and once I know it’s unique, I can start to apply labels to it. And other people can start to apply labels to it.

But then, the second play is that without additional coordination among people, I get communal value. Suddenly, when a hundred people categorize something, I’ve got the classic paralog curve. If you rank tags by frequency, you invariably get some version of a paralog curve. So you can start to answer questions like: How strong is the communal agreement around the core of what this thing is or is not about? And you can see places where there’s a strong agreement within the community—for example, that this item is definitely about software, and the other tags that have been appended to it are much more about the local context. In other cases, there’s a lot more confusion and it’s a much flatter curve. What you’re seeing here is people coming from different points of view, and the community is effectively trying to figure out whether there’s a core shared value there or not.

I don’t know if you saw the Ajax article that Jesse James Garret wrote in early 2005. He basically described an existing school of development, but he gave it a name and pointed to it. And he called it Ajax. When you look at the first three weeks of people tagging that article, they were tagging it “JavaScript,” “dynamic,” “web,” “programming,” “XML,” “asynchronous,” and so forth. And then slowly they began tagging it Ajax. And you could see that the label that he provided for it became the lens through which people were looking at the world. You could see the emergence of the conceptual handle in the middle of people trying to figure this out.

So that strikes me as really quite profound. When you give up the virtues of formal categorization, you’re giving up the right to handle things coherently and atomically. So context is the only thing that saves you from chaos.

As interesting as the tagging of URLs and other online content is, I think the tagging of the real world is going to be quite, quite remarkable.

And what’s really profound is that this kind of tagging is bridging all of this context-deriving work with all of the stuff having to do with lowered barriers to entry. More people have mobile devices
on them that can do more than make phone call, including using them for getting access to maps and marking up those maps.

On Amateur Versus Professional Classifications

The theoretical value of a classification scheme is the ability to use it to find content in this kind of uncomplicated fashion. But in fact, the amount of education you have to go through in using a card catalog is a kind of solidarity good. So a huge amount of the value of the Library of Congress system is the fact that all the catalogers and all the users agree to use the Library of Congress system. And those are sunk costs, right? At least for the current population.

So the value of rolled-up user (amateur) labeling will start to exceed the value of professional schemes first and foremost in places where there are no professional schemes. There’s no way to catalog weblog articles right now. The professional staff doesn’t exist, and if they did exist, they could only be hired at a cost that would cripple any for-profit business. But there are lots and lots of ways to extract value as a byproduct of amateur tagging of that material.

I’ll give you an example. I periodically follow an argument about web architecture called REST—Representational State Transfer. It was articulated by Roy Fielding in his dissertation, and it’s become a way of modeling resources on the web. O’Reilly had published a very good chapter on REST in a book on web services for Perl, and I convinced them to publish it as a full, free, stand-alone PDF. Just that chapter. And then I said, “Oh, I have to alert the REST community that this is out there.” And I realized that I don’t know the REST community. I don’t hang out on those mailing lists. And then I realized that I could just go to del.icio.us and tag it “REST” and that would be enough.

So here’s a case where there was no existing way for me to do it, and the existence of Del.icio.us tags made all the difference between being able to do it moderately well and not being able to do it all.

So when you have a large distributed corpus and no financial leverage or incentive to get it cataloged professionally, those are the things we are tagging. The value of tagging is going to be clearest in these cases because the competition is weakest. In places where you actually rely on a categorization system to be a kind of reifying system, like DSN4 (the diagnostic and statistical manual from the American Psychiatric Association), we’ll still turn to professionals. Those systems are there so that you and I, looking at the same set of symptoms, will both say: “Well, this is psychosis.” But it has the additional affect of making both of us be psychiatrists in a way that fits within the culture. Right? So that’s a place where the producers of the manual and the users of the manual all exist under one professional rubric. And those places are going to be least affected by tagging, because, in many ways, it’s the professional dictates rather than the classification problem that generates this scheme in the first place. Between those two extremes, a bit at a time, you’re going to see tagging encroaching on traditional classification schemes.
On the Value Proposition of Group Coordination

The overall value that I’ve seen across multiple systems is the replacement of planning with coordination. It’s what happens any time anyone gets their first cell phone. Right? They stop making exact plans and start saying: “I’ll call you when I’m done. Ring me when you’re in the neighborhood.” And the ability to coordinate reduces the required advanced planning. Cell phones, and to some degree SMS—because they’ve been relentlessly point-to-point, because of their intellectual heritage—have solved the problem in dyads. But they don’t solve it for groups.

And what we’re starting to see with some of the people who have used context awareness online is a shift in real-world group behavior. Context awareness means that groups can dispense, to some degree, with planning in favor of coordination.

So one of the big next steps is going to be a reduction in coordination costs. For example, the weblog world existed in a kind of twilight of findability for the first four or five years of its existence. It was small enough that people were relying on strategies like blogrolls and so forth. It was the same with the early days of surfing the web. You just went from place to place, hoping you would come across some value.

It was only later, in part because of Permalink and Google and in part because weblog culture became more externally focused, that weblogs emerged from this twilight of findability into the indexable world we have now. And I suspect that with context-aware stuff, we’ll see the same pattern—strategies that lower the cost of coordination.

On Group Filtering: From Problems to Desires

I had a student, Andy, whose intuition was that the Bluetooth signature of your phone was more useful as a calling card than as a network interface. Bluetooth is so crippled by the phone companies, and the security setup means that all the promise of rendezvous-style applications hasn’t materialized. But you can always detect the existence of Bluetooth devices. And since it’s a GUID [globally unique identifier], it’s a signature. That’s what led to the “Familiar Strangers” project at Intel.

But Andy’s intuition was that you should be able to tie network-available resources to a Bluetooth address. You would just use the Bluetooth address essentially as part of a URL. So if your Bluetooth address resolves to a URL, you can make public material, like a business card, available to anyone who can detect your Bluetooth device. So I could get an RSS feed tied to someone’s Bluetooth number and use it as a signal without having to use it as a device.

Andy’s next idea, which he’s working on now, is to tie the Bluetooth signature to your music profile. His idea is that there’s a jukebox in the middle of the room. And the playlist on the jukebox reads the Bluetooth address of people there and creates a kind of Last.FM-style profile. It’s like he’s using the Last.FM database to move the music mix to the aesthetic censure of the group gathered around it.

You hear stuff like that and you think: “That is the looiest idea ever.” And of course, it requires a group of people to sign up. You’d have to get a group of friends to use it. But the notion that when you drift into the room, the music sounds more like you—that kind of context awareness is not so
directed at solving any larger, long-term problem. But the space of possibility it illuminates is so interesting.

That’s the stage I think we’re at now. When people are talking about solving thorny problems with context awareness, they’re mostly talking about problems that existed in the previous iteration of technology. But when people are doing stuff like Andy’s doing, it’s really just taking the new possibilities for granted. And that’s the stuff that’s really going to rock the world.

On Group Discovery: The Space Between Personal and Global

There’s another example that may be interesting. If you go to del.icio.us and look at the tag “nptech,” you’ll find that it’s a group of people who are working together to research not-for-profit technology. They have banded together to tag information and practices for not-for-profit technology.

They’re not doing the thing most groups do—which is to say, “Oh, we need a thesaurus, or we need a glossary, or we need to agree on X, Y, and Z.” That’s generally a bad move because it’s high-cost and low-value, and it also destroys some of the different points of view within the group that you rely on to make the group viable in the first place. They went the Shibboleth route. The said: “We’re going to pick one word that signifies membership in this group. And that word is ‘nptech.’” And when you’re tagging something relevant to the group, tag it as you would ordinarily, but add this tag.

And that becomes a kind of handle for the group effort. It’s not really a tag in the same way. It’s a hack. It labels something different. But playing with group categorization and group classification of this shared material is going to become fantastically valuable. That is my guess, in part, because the Internet is natively a group-forming medium, and, in part because we keep walking through the UNIX permissions in an incomplete way. Right? We start with a personal technology and then we make it global. Or we start with a global technology and then we make it personal. And we only back into group—someplace between user and world—after the fact.
Look at del.icio.us right now. There’s an individual del.icio.us user and there’s the front page, which is everybody rolled-up altogether. But now we’re starting to see people press del.icio.us into service for: “This is what my research group is looking at.” Which is a different kind of value than knowing what an individual user is looking at, and it’s also a different kind of value from what everyone is looking at.

Tags were not invented to declare group boundaries, of making them visible, and then making it possible to participate in the group. Right? There are both intrinsic and extrinsic group boundaries, right? That’s a declared group boundary. But you can absolutely bet that the tag REST pulls together a group of programmers who have a bunch of other similar bits of worldview and who are largely known to each other. The one different thing is that they don’t think of themselves as being members of the group that they participate in. So it’s sort of high- and low-edge conditions. But all these groups are discoverable through tags. They’re kind of epistemological communities that are discoverable by going over the corpus after the fact and looking for similar tags.

It turns out that the idea of membership has been made so lightweight that you suddenly start getting the kind of value that was previously possible only in an institutional context. You can now get that kind of value from much more loosely aggregated groups of individuals.