O'REILLY®

Ambient Computing

How Invisible Hardware, Self-Starting Apps, and Nonstop Surveillance Reshapes Our Public and Private Lives

Mike Barlow

Ambient Computing

How Invisible Hardware, Self-Starting Apps, and Nonstop Surveillance Reshape our Public and Private Lives

Mike Barlow

Beijing • Boston • Farnham • Sebastopol • Tokyo O'REILLY®

Ambient Computing
by Mike Barlow

Printed in the United States of America.

Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472.

O'Reilly books may be purchased for educational, business, or sales promotional use. Online editions are also available for most titles (*http://safaribooksonline.com*). For more information, contact our corporate/institutional sales department: 800-998-9938 or *corporate@oreilly.com*.

Editor: Susan Conant
Production Editor: Nicholas Adams
Interior Designer: David Futato
Cover Designer: Randy Comer
Illustrator: Rebecca Demarest

June 2016: First Edition

Revision History for the First Edition
2016-06-17: First Release

978-1-491-95482-9

[LSI]

Table of Contents

Ambient Computing

A medley of pervasive networks and interoperable information technologies has produced a new state of persistent surveillance, connectivity, and analysis. No more sitting down at laptops or reaching for mobile devices.

We've entered the Age of Ambient Computing, a convergence of cheap sensors, wireless connectivity, increasingly powerful microchips, and advanced analytics that is redefining the nature of public and private spaces in the world's developed economies.

The good news: ambient computing is largely hands-free and effortless, at least from the perspective of the average citizen. The bad news: it's like the weather—good or bad, you can't avoid it.

Ready or not, we're surrounded by sensors that collect data and send it via networks to analytics that convert the data into information for a potpourri of end users ranging from corporate marketers to government security agencies.

Here are some hypothetical examples: you walk into a conference hall and a Bluetooth Low Energy (BLE) beacon beams a short coded message announcing its proximity. Assuming that you have an app on your phone that can interpret and identify the beacon, you see a welcome message on your lock screen, along with directions to the room where a session you've signed up to attend is starting in five minutes. Your phone also displays information about when lunch will be served, and confirms your choice of the chicken, fish, or vegan entree.

That's a fairly benign and innocuous case. Let's go a little deeper and assume the conference hall's ambient computing systems can also glean bits of personal data from their interactions with your smart phone. Perhaps you will receive an allergy warning based on your own personal health profile or on the health profiles of similar people in your age cohort.

Or perhaps the system will identify you as a VIP and someone will offer you a free glass of champagne. On a darker note, the system might identify you as a security threat, and you might be detained or arrested.

Let's say you're running in Central Park and your smart phone detects a potentially dangerous arrhythmia in your heartbeat. It might launch an app that gives you an onscreen alert and suggests that you slow down. Or it might skip the alert and send for an ambulance.

Enabling those types of scenario will require lots of work behind the scenes. Issues of network connectivity, device security, and personal privacy must be uncovered, discussed, and resolved. Companies will need to develop practical solutions that are easy to use and can be scaled for a variety of markets and situations.

All of that suggests the need for a rough guide or framework to help developers, investors, vendors, purchasers, and users get a better idea of the components and capabilities required for ambient computing scenarios.

Some Assembly Required

The easiest way to begin developing a framework is by imagining the layers of an ambient computing technology stack (see Figure 1-1). The bottom layer includes sensors, devices, and other data sources. The next layer provides network connectivity and data transport.

The analytics layer is next. This is where the heavy lifting starts, and where the process of converting raw data into useful information begins. Above the analytics layer sits the user experience layer, which includes applications, tools, and interfaces.

The stack is a work in progress and is continuously evolving. Depending on your perspective, the stack might have more layers, fewer layers, or added dimensions.

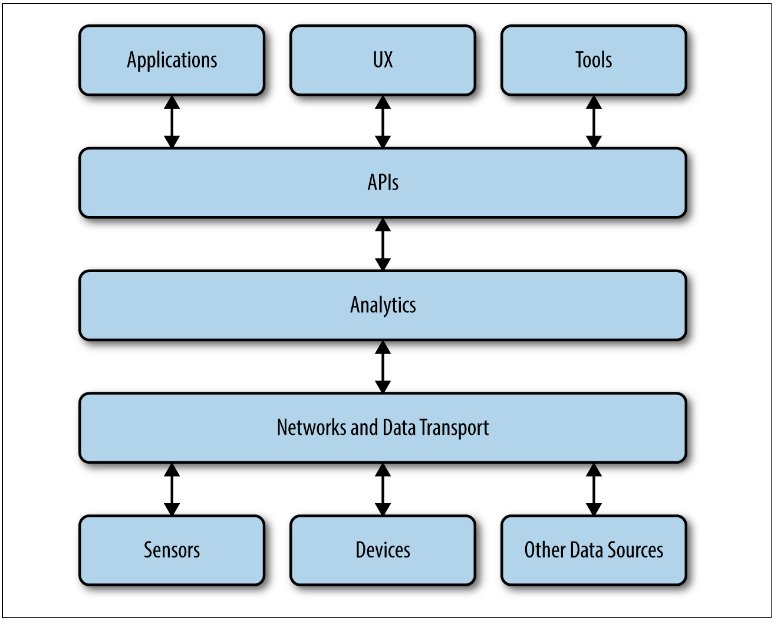

Figure 1-1. The ambient computing stack

Matthew Gast, for instance, predicts the need for an API layer. Gast is the director of software product management at Aerohive Networks. He is active in the WiFi community and has served as a leader on several industry standards committees, including as chair of the current revision of the 802.11 standard. Gast has written extensively about WiFi, and he is the author of several O'Reilly books.

From his viewpoint, the API layer would sit immediately below the applications layer. "APIs enable programmers to build context that represents what is important to users," he writes in "Searching for the software stack for the physical world," an O'Reilly article. Mashing up data from multiple sources to create useful information is a key part of the ambient computing process.

For example, monitoring "room temperature" can depend on collecting and combining data from multiple sensors placed around the room to gather data on temperature, humidity, sunlight, and other physical phenomena. Managing the climate of an individual room can require orchestrating the operations of heating, cooling, air circulation, and lighting systems.

"The real-world importance of data is apparent only when it's combined with other data," says Gast. In other words, context is absolutely critical. APIs allow programmers to create practical bridges between the higher and lower layers of the ambient computing stack.

Real-Time Analytics Enable Real-Time Decision Making

In a sense, the ambient computing stack is a context-creation engine that ingests data, analyzes it, and then presents information that can be used by humans—or by autonomous systems—for making contextually relevant decisions in real time.

The rise of data science has given us a fresh perspective on what's doable with analytics. With an assist from Moore's Law, we're moving those analytics from the lab to the real world with rapid speed. Hardware is no longer the primary constraint; now the challenge is writing better and better software for our analytics.

Xiaofan (Fred) Jiang is an assistant professor in the Department of Electrical Engineering at Columbia University. He also works with the university's Data Science Institute and its Intelligent and Connected Systems Lab (ICSL).

Jiang and his colleagues are developing low-cost, accurate, and scalable solutions for monitoring air quality in cities like Beijing, where the sheer size of the city creates huge problems for anyone trying to collect data.

One of the problems, he explains, is that cities often focus on buying expensive hardware, instead of looking for cheaper and more practical alternatives. "Air quality is a global problem that impacts billions of people," says Jiang. "Our idea is combining cheap sensors, software, analytics, and the cloud to improve data quality. From better data, we create knowledge and actionable feedback to improve everyday life." (See Figure 1-2.)

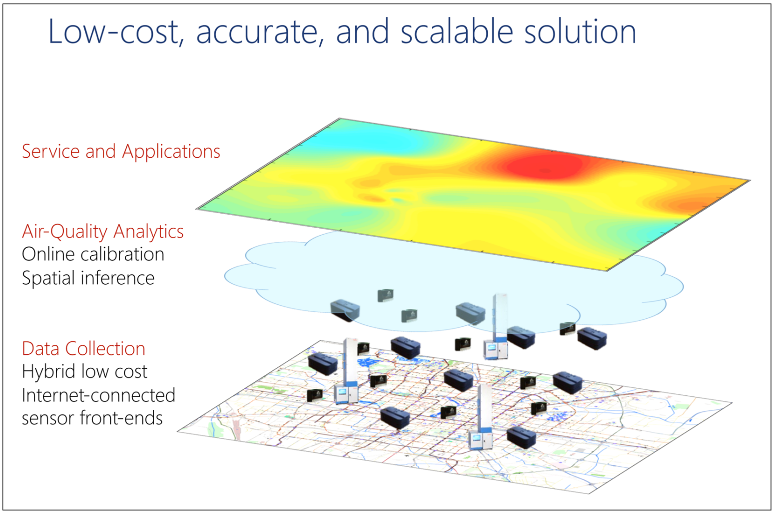

Figure 1-2. The layers of a practical air quality monitoring system. Image credit: Professor Xiaofan (Fred) Jiang, Department of Electrical Engineering, Columbia University

Essentially, Jiang's team is trading small numbers of highly accurate and expensive sensors for large numbers of inexpensive and less accurate sensors. It's like having a swarm of sensors.

"The quality of our sensors isn't very good, but we connect them to the cloud through cellular and Bluetooth, and improve data quality computationally. We have a stationary version and a portable version so people can take sensors with them as they travel around the city, which gives us more spatial coverage," he explains.

"Even though each sensor is less accurate, having lots of them gives us a much broader set of data and enables us to continuously recalibrate them. By connecting them to the cloud, we can bring all the information together, apply various analytics, and improve the overall accuracy of the system," says Jiang.

Improved accuracy—at lower costs—translates into better information for government agencies that monitor sources of pollution. "Now we can quickly generate accurate maps showing where the pollution comes from," says Jiang, as shown in Figure 1-3.

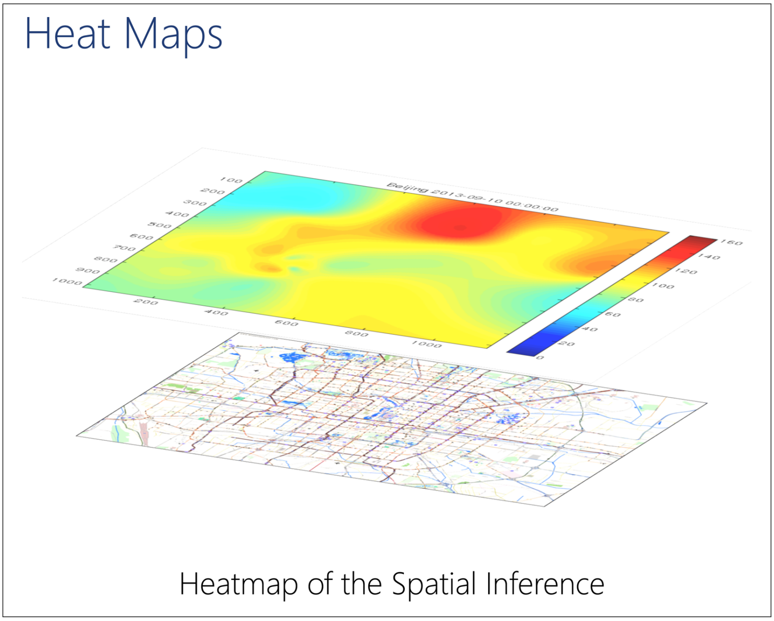

Figure 1-3. Real-time heat maps generated by the air quality monitoring system can pinpoint sources of pollution. Image credit: Professor Xiaofan (Fred) Jiang, Department of Electrical Engineering, Columbia University

Citizen Sensing and the TMZ Lifestyle

Jiang uses the term "participatory sensing" to describe a fundamental aspect of the air quality monitoring scenario.

Imagine how much useful data you would collect by putting sensors in every taxicab or on every bicycle in a city. It might sound like Big Brother, until you remember that most of us already carry dozens of sensors around with us in our smart phones.

Even in the rare moments when we aren't holding or carrying our smart phones, we're surrounded by cameras, microphones, and other devices capable of capturing, recording, and relaying data about our activities.

"We live in an era of ubiquitous surveillance, but the surveillance isn't coherent," says Ari Gesher, director of software engineering at Kairos Aerospace. Instead of just one Big Brother or Peeping Tom looking over our shoulders or gazing through our curtains, there are legions of organizations, agencies, and individuals continually spying on us. We even spy on ourselves and share intimate details of our lives on social networks.

Nonstop surveillance has led to what Gesher describes as the "TMZ lifestyle," a reference to the popular website devoted to celebrity gossip. "We're entering a future in which we're all wearing devices that can record everything that's going on around us," he says. "Everyone will have the capability of posting 'gotcha' photos of everyone else."

As a result, he suggests, people will become fearful of interacting with other people. They will avoid entering public spaces. They will spend countless hours devising tactics for avoiding surveillance by employers, police, family, and even friends. "It's a nightmare scenario. Very dystopian," he says.

Alternatively, society could develop "polite privacy" frameworks that would reduce or mitigate the risks of ubiquitous surveillance and sensing. Although voluntary and unenforceable, such frameworks would at least blunt the worst aspects of the TMZ ethos.

A privacy framework would enable you to set your own personal privacy parameters, similar to the way you use the privacy settings on Facebook. For example, you might have a group of friends or colleagues who are allowed to share your photos, status updates, and location information. You might have another group that's allowed to share your photos and status updates, but not your location information. And you might have another group that can share your photos, but not your status updates and location information.

Your privacy profile would probably reside somewhere in the cloud, but would be accessible to companies, organizations, and other individuals via your smart phone or mobile device. As you walk or travel around, various beacons would be constantly querying your phone or tablet and determining your privacy parameters and deciding whether or not you're fair game. Ideally, the privacy frameworks would provide seamless and effortless transmission of your preferences.

"It's not an iron-clad solution, but it gives us some assurance," says Gesher. "At least we'll know that when our images or information about us is being recorded, our preferences for sharing those images or information are being recorded as well."

Gesher does not expect companies like Google, Facebook, or Twitter to develop such frameworks. “There’s no compelling business reason and no proven market,” he says. But he hopes that “someone at a university or within the research community” will create an open source program or protocol that would get the ball rolling.

He’s reasonably confident that research would eventually lead to practical privacy frameworks that could be used widely. The alternative would be waiting for the government to mandate comprehensive data privacy regulations for ambient computing, which seems unlikely.

“We’re running out of space to live our private lives,” says Gesher. “We live in a panopticon and we’ve become our own jailers.”

Emotional Analytics

So far, we’ve been talking about ambient computing systems that feed off information that’s more or less voluntarily shared. In some cases, our consent is explicit; in other cases it’s implied. For example, unless you turn off the WiFi in your smart phone, you automatically announce your presence when you walk past a rows of restaurants or park your car at the shopping mall.

But there’s also a whole spectrum of data generated by processes we have little control over. Our voices and the way we speak are windows into our emotional states. Audio sensors can easily pick up nuances in speech and relay data to analytics that can tell with astonishing accuracy whether you’re happy, sad, angry, fearful, or disgusted.

Julia Hirschberg is Percy K. and Vida L. W. Hudson Professor of Computer Science and Chair of the Computer Science Department at Columbia University. She worked at Bell Laboratories and AT&T Laboratories-Research from 1985–2003, where she created the Human-Computer Interface Research Department.

Hirschberg and her colleagues specialize in analyzing emotions in spoken language. That kind of analysis is especially important for organizations that operate large customer contact centers. If you’re running a contact center, you want to automate as much of the call-handling processes as possible.

A contact center's automated systems can handle a wide range of problems from callers, but some situations still require the skills of a human operator. The hard part is determining which calls require human intervention or immediate attention. That's where the capability of automatically distinguishing between impatience, frustration, anger, and boiling rage comes in handy.

Software can be taught to recognize emotional states by analyzing changes in pitch, speed, and volume when people are talking. Emotional speech analysis is consistently more accurate than human judgment or lie detector machines.

The Walls Have Ears

But the potential for applying emotional state analysis ranges far beyond customer contact centers. Hirschberg sees a role for emotional analytics in hospitals, geriatric facilities, smart homes, offices, schools, and public spaces such as train stations, airports, and sports stadiums.

"Corporations might want to monitor levels of employee contentment, schools might want to check for the presence of extreme levels of anger and frustration to detect bullying and conflict. In public places, monitoring the emotional state of the crowd or of individuals might be a useful way of identifying potential threats of violence," Hirschberg says.

It seems reasonable to imagine that emotional state analyzers will become standard features of ambient computing environments. Would the ubiquity of emotional state analytics redefine the nature of privacy in public spaces?

"There are certainly some situations in which you would want to opt out of such monitoring," says Hirschberg. On the other hand, emotional state monitoring in public spaces might become a sort of necessary evil for protecting society from violence.

When you enter a public space, you would also enter into an implicit agreement with the organization that owns or manages the space; essentially, you would yield a portion of your personal privacy and "opt in" to being covertly monitored while you're in the space.

It's All About Creating Context

F. Mark Modzelewski is chief of staff at Estimote, a tech company that makes beacons. From his point of view, smart phones are actually "blind phones." Your smart phone knows exactly where it is in terms of location data, but it doesn't know where it is contextually.

For example, your smart phone knows that it's on the corner of Lexington Avenue and 40th Street. But it doesn't know that you're hungry and trying to decide which one of a dozen nearby restaurants is the best choice for a quick lunch. If the restaurants had beacons, the phone could interact with them and flash their menus on your lock screen.

"It's not the phones and the beacons that really matter, it's the context they create," says Modzelewski. The combination of the phone and the beacon merely set the stage for an ambient computing scenario. Making the connection is easy; what's harder is establishing the context for a continuous series of interactions.

In 2014, Estimote introduced "sticker beacons" with a built-in accelerometer and temperature sensors. Placing a sticker beacon on your bike makes it easy to track all of your rides with your smart phone. It also makes it harder for thieves to abscond with your bicycle, because you can track its location with your smart phone.

A shoe retailer, for example, might place tiny sticker beacons on the soles of running shoes on display. When a prospective customer picks up a shoe, the beacon would launch an instructional video, complete with detailed specs about the shoe and helpful running tips, on a screen near the customer.

The ability to sense "micro location" is also a key to the future of ambient computing, says Modzelewski. When you walk past the dairy aisle at the supermarket, for instance, a beacon would launch the shopping list app on your phone and remind you to pick up a quart of milk or a container of yogurt.

The Newer Normal

Like many tech execs, Modzelewski expresses irritation over terms such as "Internet of Things," "Industrial Internet," and "Internet of

Everything." Soon, it will be uncommon to find a product or device that isn't connected to the Internet.

"We do not need a mental crutch to visualize the connected world anymore. We see it everywhere, because the Internet is ubiquitous. Connectivity is the new norm. It's part of the fabric of the physical world and literally has engulfed us," he writes in a recent post on LinkedIn. "Pointing out difference between the Internet of Things and just the Internet is totally pointless."

SIGFOX, a global communications provider headquartered in France, is pushing to make connectivity a nonissue, by offering low-power, low-cost solutions designed to support long-term, large-scale sensor deployments in places like forests and farmlands.

Having thousands of sensors placed on the ground gives farmers better insights about when crops need more water. Sticking heat sensors on trees gives fire fighters a chance to control small woodland blazes before they morph into destructive raging infernos.

From the perspective of SIGFOX, bringing ambient computing to large areas requires armies of low-energy devices operating within in ultra-narrow bands of the radio spectrum. That's a different approach than the scenarios envisioned by major carriers such as Verizon and AT&T, which would be happy to see ambient computing systems running on their existing 3G and 4G networks.

"3G and 4G are beautiful when you're streaming HD video, but they're overkill by 1000x for most of the use cases we envision. It's like renting a stretch limo to drive your kids to the corner convenience store for ice cream," says Allen Proithis, president of SIGFOX in North America.

In addition to preventing forest fires or detecting drought in rural areas, low-power sensors and low-cost networks are essential for smart cities. "When you put sensors in garbage cans, you know when they're full. A smart city can save 30 to 40 percent on garbage collection costs by only picking up full cans," says Proithis.

Eventually, he believes, ambient computing solutions will rely on networks that are essential blends of "complementary connections." The most likely scenarios would have devices connecting to systems over multiple networks, including cellular, WiFi, Bluetooth, ZigBee, and other forms of wireless communication. In other words, different types of connectivity will he handled by different types of networks.

Smarter, Faster, and More Powerful

The progress of ambient computing will depend to a large degree on the evolution of microchip technology. In February 2016, MIT researchers unveiled a new and radically more powerful chip designed specifically for neural networks. The new chip is called "Eyeriss," a loose acronym for "energy-efficient reconfigurable accelerator for deep convolutional neural networks," and it represents a critical leap forward in the development of ambient computing.

Essentially, Eyeriss is so efficient and powerful that it enables individual nodes or devices to process and manipulate data at speeds that now require multiple nodes working in parallel. With Eyeriss, less information needs to flow across the network. Results are produced faster, and perhaps most important, privacy is preserved, since most of the information required to perform a computational task never has to leave the device.

"The idea is that sensors can do more jobs independently," says Yu-Hsin Chen, a member of the MIT research team that designed Eyeriss. Instead of exchanging raw data across insecure networks, sensors equipped with Eyeriss keep data "at home" and only share the results of their labor.

Upcoming generations of connected cars, trains, planes, ships, and surgical robots will rely on chips like Eyeriss to perform life-or-death calculations at lightning speed, while preserving some degrees of privacy and security.

The new chips will also pave the way for integrating convolutional neural networks into ambient computing scenarios. Since convolutional neural networks are like A.I. on steroids, they would enable the development of ambient computing systems that are capable of learning and evolving.

In the not-too-distant future, ambient computing environments will do more than merely recognize you, understand your preferences, and analyze your emotional state. They will be capable of changing and adapting to suit your personal taste and mood.

That, of course, is the upside of ambient computing. The downside is that those same systems will also know your FICO score, marital status, political affiliation, race, religion, and country of origin. It's not hard to imagine scenarios in which that information could be used to create environments that are unwelcoming or hostile to some groups of people.

It's clear that we've moved beyond the "Internet of Things" and are heading down a path to ubiquitous connectivity and computing. We can only hope that our social, legal, and moral frameworks will evolve as well.

About the Author

Mike Barlow is an award-winning journalist, author, and commentator. He is the author of *Learning to Love Data Science* (O'Reilly Media, 2015), and the coauthor of *The Executive's Guide to Enterprise Social Media Strategy* (Wiley, 2011), and *Partnering with the CIO: The Future of IT Sales Seen Through the Eyes of Key Decision Makers* (Wiley, 2007). He is also the writer of many articles, reports, and white papers on numerous topics such as smart cities, social networking, cloud computing, IT infrastructure, predictive maintenance, data analytics, and data visualization.

Over the course of a long career, Barlow was a reporter and editor at several respected suburban daily newspapers, including *The Journal News* and the *Stamford Advocate*. His feature stories and columns appeared regularly in *The Los Angeles Times*, *Chicago Tribune*, *Miami Herald*, *Newsday* and other major US dailies. He has also written extensively for O'Reilly Media.

A graduate of Hamilton College, he is a licensed private pilot, avid reader, and enthusiastic ice hockey fan.