Web^2

Special Report

Web Squared: Web 2.0 Five Years On

By Tim O'Reilly and John Battelle

web2.0 SUMMIT

CO-PRODUCED BY O'REILLY® techweb™ A Division of United Business Media LLC

Contents

Disclosure: *Via our venture investing arm, O'Reilly AlphaTech Ventures, O'Reilly is an investor in three of the companies mentioned in this article: Wesabe, AMEE, and bit.ly*

Web Squared: Web 2.0 Five Years On

Five years ago, we launched a conference based on a simple idea, and that idea grew into a movement. The original Web 2.0 Conference (now the Web 2.0 Summit) was designed to restore confidence in an industry that had lost its way after the dotcom bust. The Web was far from done, we argued. In fact, it was on its way to becoming a robust platform for a culture-changing generation of computer applications and services.

In our first program, we asked why some companies survived the dotcom bust, while others had failed so miserably. We also studied a burgeoning group of startups and asked why they were growing so quickly. The answers helped us understand the rules of business on this new platform.

Chief among our insights was that "the network as platform" means far more than just offering old applications via the network ("software as a service"); it means building applications that literally get better the more people use them, harnessing network effects not only to acquire users, but also to learn from them and build on their contributions.

From Google and Amazon to Wikipedia, eBay, and craigslist, we saw that the value was facilitated by the software, but was co-created by and for the community of connected users. Since then, powerful new platforms like YouTube, Facebook, and Twitter have demonstrated that same insight in new ways. Web 2.0 is all about harnessing collective intelligence.

Collective intelligence applications depend on managing, understanding, and responding to massive amounts of user-generated data in real time. The "subsystems" of the emerging internet operating system are increasingly data subsystems: location, identity (of people, products, and places), and the skeins of meaning that tie them together. This leads to new levers of competitive advantage: Data is the "Intel Inside" of the next generation of computer applications.

Today, we realize that these insights were not only directionally right, but are being applied in areas we only imagined in 2004. The smartphone revolution has moved the Web from our desks to our pockets. Collective intelligence applications are no longer being driven solely by humans typing on keyboards but, increasingly, by sensors. Our phones and cameras are being turned into eyes and ears for applications; motion and location sensors tell where we are, what we're looking at, and how fast we're moving. Data is being collected, presented, and acted upon in real time. The scale of participation has increased by orders of magnitude.

With more users and sensors feeding more applications and platforms, developers are able to tackle serious real-world problems. As a result, the Web opportunity is no longer growing arithmetically; it's growing exponentially. Hence our theme for this year: Web Squared. 1990–2004 was the match being struck; 2005–2009 was the fuse; and 2010 will be the explosion.

Ever since we first introduced the term "Web 2.0," people have been asking, "What's next?" Assuming that Web 2.0 was meant to be a kind of software version number (rather than a statement about the second coming of the Web after the dotcom bust),

we're constantly asked about "Web 3.0." Is it the semantic web? The sentient web? Is it the social web? The mobile web? Is it some form of virtual reality?

It is all of those, and more.

The Web is no longer a collection of static pages of HTML that describe something in the world. Increasingly, the Web is the world—everything and everyone in the world casts an "information shadow," an aura of data which, when captured and processed intelligently, offers extraordinary opportunity and mind-bending implications. Web Squared is our way of exploring this phenomenon and giving it a name.

Redefining Collective Intelligence: New Sensory Input

To understand where the Web is going, it helps to return to one of the fundamental ideas underlying Web 2.0, namely that successful network applications are systems for harnessing collective intelligence.

Many people now understand this idea in the sense of "crowdsourcing," meaning that a large group of people can create a collective work whose value far exceeds that provided by any of the individual participants. The Web as a whole is a marvel of crowdsourcing, as are marketplaces such as those on eBay and craigslist, mixed media collections such as YouTube and Flickr, and the vast personal lifestream collections on Twitter, MySpace, and Facebook.

Many people also understand that applications can be constructed in such a way as to direct their users to perform specific tasks, like building an online encyclopedia (Wikipedia), annotating an online catalog (Amazon), adding data points onto a map (the many web mapping applications), or finding the most popular news stories (Digg, Twine). Amazon's Mechanical Turk has gone so far as to provide a generalized platform for harnessing people to do tasks that are difficult for computers to perform on their own.

But is this really what we mean by collective intelligence? Isn't one definition of intelligence, after all, that characteristic that allows an organism to learn from and respond to its environment? (Please note that we're leaving aside entirely the question of self-awareness. For now, anyway.)

Imagine the Web (broadly defined as the network of all connected devices and applications, not just the PC-based application formally known as the World Wide Web) as a newborn baby. She sees, but at first she can't focus. She can feel, but she has no idea of size till she puts something in her mouth. She hears the words of her smiling parents, but she can't understand them. She is awash in sensations, few of which she understands. She has little or no control over her environment.

Gradually, the world begins to make sense. The baby coordinates the input from multiple senses, filters signal from noise, learns new skills, and once-difficult tasks become automatic.

The question before us is this: Is the Web getting smarter as it grows up?

Consider search—currently the lingua franca of the Web. The first search engines, starting with Brian Pinkerton's webcrawler, put everything in their mouth, so to speak. They hungrily followed links, consuming everything they found. Ranking was by brute force keyword matching.

In 1998, Larry Page and Sergey Brin had a breakthrough, realizing that links were not merely a way of finding new content, but of ranking it and connecting it to a more sophisticated natural language grammar. In essence, every link became a vote, and votes from knowledgeable people (as measured by the number and quality of people who in turn vote for them) count more than others.

Modern search engines now use complex algorithms and hundreds of different ranking criteria to produce their results. Among the data sources is the feedback loop generated by the frequency of search terms, the number of user clicks on search results, and our own personal search and browsing history. For example, if a majority of users start clicking on the fifth item on a particular search results page more often than the first, Google's algorithms take this as a signal that the fifth result may well be better than the first, and eventually adjust the results accordingly.

Now consider an even more current search application, the Google Mobile App for the iPhone. The application detects the movement of the phone to your ear, and automatically goes into speech recognition mode. It uses its microphone to listen to your voice, and decodes what you are saying by referencing not only its speech recognition database and algorithms, but also the correlation to the most frequent search terms in its search database. The phone uses GPS or cell-tower triangulation to detect its location, and uses that information as well. A search for "pizza" returns the result you most likely want: the name, location, and contact information for the three nearest pizza restaurants.

All of a sudden, we're not using search via a keyboard and a stilted search grammar, we're talking to and with the Web. It's getting smart enough to understand some things (such as where we are) without us having to tell it explicitly. And that's just the beginning.

And while some of the databases referenced by the application—such as the mapping of GPS coordinates to addresses—are "taught" to the application, others, such as the recognition of speech, are "learned" by processing large, crowdsourced data sets.

Clearly, this is a "smarter" system than what we saw even a few years ago. Coordinating speech recognition and search, search results and location, is similar to the "hand-eye" coordination the baby gradually acquires. The Web is growing up, and we are all its collective parents.

Cooperating Data Subsystems

In our original Web 2.0 analysis, we posited that the future "internet operating system" would consist of a series of interoperating data subsystems. The Google Mobile Application provides one example of how such a data-driven operating system might work.

In this case, all of the data subsystems are owned by one vendor—Google. In other cases, as with Apple's iPhoto '09, which integrates Flickr and Google Maps as well as Apple's own cloud services, an application uses cloud database services from multiple vendors.

As we first noted back in 2003, data is the "Intel Inside" of the next generation of computer applications. That is, if a company has control over a unique source of data that is required for applications to function, they will be able to extract monopoly rents from the use of that data. In particular, if a database is generated by user contribution, market leaders will see increasing returns as the size and value of their database grows more quickly than that of any new entrants.

We see the era of Web 2.0, therefore, as a race to acquire and control data assets. Some of these assets—the critical mass of seller listings on eBay, or the critical mass of classified advertising on craigslist—are application-specific. But others have already taken on the characteristic of fundamental system services.

Take for example the domain registries of the DNS, which are a backbone service of the Internet. Or consider CDDB, used by virtually every music application to look up the metadata for songs and albums. Mapping data from providers like Navteq and TeleAtlas is used by virtually all online mapping applications.

There is a race on right now to own the social graph. But we must ask whether this service is so fundamental that it needs to be open to all.

It's easy to forget that only 15 years ago, email was as fragmented as social networking is today, with hundreds of incompatible email systems joined by fragile and congested gateways. One of those systems—internet RFC 822 email—became the gold standard for interchange.

We expect to see similar standardization in key internet utilities and subsystems. Vendors who are competing with a winner-takes-all mindset would be advised to join together to enable systems built from the best-of-breed data subsystems of cooperating companies.

How the Web Learns: Explicit vs. Implicit Meaning

But how does the Web learn? Some people imagine that for computer programs to understand and react to meaning, meaning needs to be encoded in some special taxonomy. What we see in practice is that meaning is learned "inferentially" from a body of data.

Speech recognition and computer vision are both excellent examples of this kind of machine learning. But it's important to realize that machine learning techniques apply to far more than just sensor data. For example, Google's ad auction is a learning system, in which optimal ad placement and pricing is generated in real time by machine learning algorithms.

In other cases, meaning is "taught" to the computer. That is, the application is given a mapping between one structured data set and another. For example, the association between street addresses and GPS coordinates is taught rather than learned. Both data sets are structured, but need a gateway to connect them.

It's also possible to give structure to what appears to be unstructured data by teaching an application how to recognize the connection between the two. For example, You R Here, an iPhone app, neatly combines these two approaches. You use your iPhone camera to take a photo of a map that contains details not found on generic mapping applications such as Google maps—say a trailhead map in a park, or another hiking map. Use the phone's GPS to set your current location on the map. Walk a distance away, and set a second point. Now your iPhone can track your position on that custom map image as easily as it can on Google maps.

Some of the most fundamental and useful services on the Web have been constructed in this way, by recognizing and then teaching the overlooked regularity of what at first appears to be unstructured data.

Ti Kan, Steve Scherf, and Graham Toal, the creators of CDDB, realized that the sequence of track lengths on a CD formed a unique signature that could be correlated with artist, album, and song names. Larry Page and Sergey Brin realized that a link is a vote. Marc Hedlund at Wesabe realized that every credit card swipe is also a vote, that there is hidden meaning in repeated visits to the same merchant. Mark Zuckerberg at Facebook realized that friend relationships online actually constitute a generalized social graph. They thus turn what at first appeared to be unstructured into structured data. And all of them used both machines and humans to do it.

Key takeaway: *A key competency of the Web 2.0 era is discovering implied metadata, and then building a database to capture that metadata and/or foster an ecosystem around it.*

Web Meets World: The "Information Shadow" and the "Internet of Things"

Say "sensor-based applications," and many people might imagine a world of applications driven by RFID tags or ZigBee modules. This future is conveniently far off, with test deployments and a few exciting early stage applications. But what many people fail to notice is how far along the sensor revolution already is. It's the hidden face of the mobile market, and its most explosive opportunity.

Today's smartphones contain microphones, cameras, motion sensors, proximity sensors, and location sensors (GPS, cell-tower triangulation, and even in some cases, a compass). These sensors have revolutionized the user interface of standalone applications—you have only to play with Smule's Ocarina for the iPhone to see that.

But remember: mobile applications are connected applications. The fundamental lessons of Web 2.0 apply to any network application, whether web- or mobile phone-based (and the lines between the two are increasingly blurred). Sensor-based applications can be designed to get better the more people use them, collecting data that creates a virtuous feedback loop that creates more usage. Speech recognition in Google Mobile App is one such application. New internet-connected GPS applications also have built-in feedback loops, reporting your speed and using it to estimate arrival time based on its knowledge of traffic ahead of you. Today, traffic patterns are largely estimated; increasingly, they will be measured in real time.

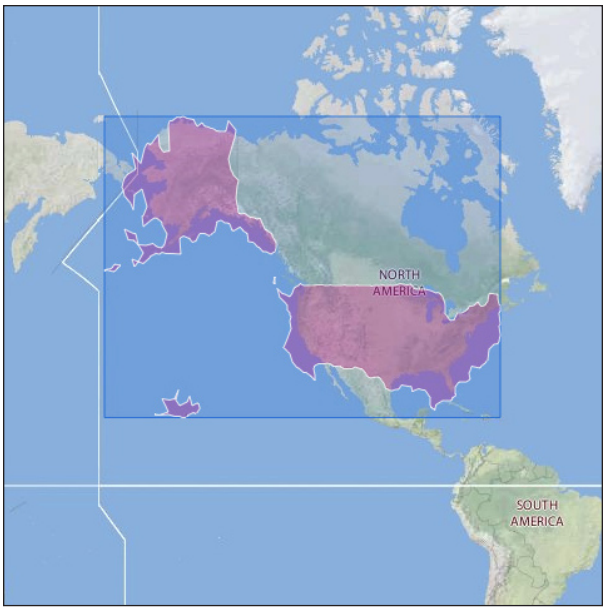
The Net is getting smarter faster than you might think. Consider geotagging of photos. Initially, users taught their computers the association between photos and locations by tagging them. When cameras know where they are, every photo will be geotagged, with far greater precision than the humans are likely to provide.

And the increased precision in one data set increases the potential of another. Consider the accuracy of the maps shown below generated by geotagged Flickr photos.

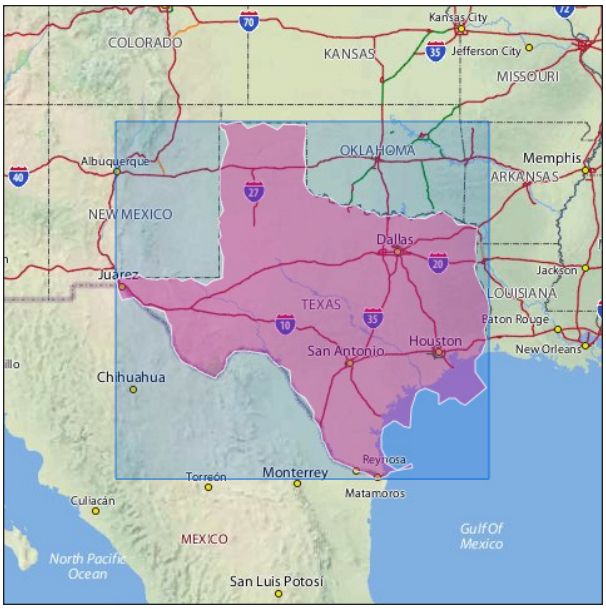
How much more accurate will these maps be when there are billions of photos?

Nor will the training wheels for the Net's visual sensor network be limited to location.

It's still in its early days, but the face recognition in Apple iPhoto '09 is pretty good. At what point are enough faces tagged with names that the system is able to show you only the people it doesn't recognize? (Whether or not Apple imagines providing this data as a system service is an open question; whether someone else does it as a network service is assuredly not.)



Flickr geotag map of USA
http://flickr.com/photos/straup/2972130238/



Flickr geotag map of Texas
http://flickr.com/photos/straup/2971287541/

The Wikitude travel guide application for Android takes image recognition even further. Point the phone's camera at a monument or other point of interest, and the application looks up what it sees in its online database (answering the question, "What looks like that somewhere around here?") The screen shows you what the camera sees, so it's like a window but with a heads-up display of additional information about what you're looking at. It's the first taste of an "augmented reality" future. It superimposes distances to points of interest, using the compass to keep track of where you're looking. You can sweep the phone around and scan the area for nearby interesting things.

Layar takes this idea even further, promising a framework for multiple layers of augmented reality content accessed through the camera of your mobile phone.

Think of sensor-based applications as giving you superpowers. Darkslide gives you super eyesight, showing you photos near you. iPhone Twitter apps can "find recent tweets near you" so you can get super hearing and pick up the conversations going on around you.

Photosynth, Gigapixel Photography, and Infinite Images

The increasing richness of both sensor data and machine learning will lead to new frontiers in creative expression and imaginative reconstruction of the world.

Microsoft's Photosynth demonstrates the power of the computer to synthesize 3D images from crowdsourced photographs. Gigapixel photography reveals details that were invisible even to people on the scene. Adobe's Infinite Images reveals something even more startling: the ability of the computer to synthesize imaginary worlds that never existed, extrapolating a complete 3D experience from a set of photos. The video demonstration needs to be seen to be believed:

All of these breakthroughs are reflections of the fact noted by Mike Kuniavsky of ThingM, that real world objects have "information shadows" in cyberspace. For instance, a book has information shadows on Amazon, on Google Book Search, on Goodreads, Shelfari, and LibraryThing, on eBay and on BookMooch, on Twitter, and in a thousand blogs.

A song has information shadows on iTunes, on Amazon, on Rhapsody, on MySpace, on Facebook. A person has information shadows in a host of emails, instant messages, phone calls, tweets, blog postings, photographs, videos, and government documents. A product on the supermarket shelf, a car on a dealer's lot, a pallet of newly mined boron sitting on a loading dock, a storefront on a small town's main street—all have information shadows now.

In many cases, these information shadows are linked with their real world analogues by unique identifiers: an ISBN or ASIN, a part number, or getting more individual, a social security number, a vehicle identification number, or a serial number. Other identifiers are looser, but identity can be triangulated: a name plus an address or phone number, a name plus a photograph, a phone call from a particular location undermining what once would have been a rock-solid alibi.

Many who talk about the Internet of Things assume that what will get us there is the combination of ultra-cheap RFID and IP addresses for everyday objects. The assumption is that every object must have a unique identifier for the Internet of Things to work.

What the Web 2.0 sensibility tells us is that we'll get to the Internet of Things via a hodgepodge of sensor data contributing, bottom-up, to machine-learning applications that gradually make more and more sense of the data that is handed to them. A bottle of wine on your supermarket shelf (or any other object) needn't have an RFID tag to join the Internet of Things, it simply needs you to take a picture of its label. Your mobile phone, image recognition, search, and the sentient web will do the rest. We don't have to wait until each item in the supermarket has a unique machine-readable ID. Instead, we can make do with bar codes, tags on photos, and other "hacks" that are simply ways of brute-forcing identity out of reality.

There's a fascinating fact noted by Jeff Jonas in his work on identity resolution. Jonas' work included building a database of known US persons from various sources. His database grew to about 630 million "identities" before the system had enough information to identify all the variations. But at a certain point, his database began to learn, and then to shrink. Each new load of data made the database smaller, not bigger. 630 million plus 30 million became 600 million, as the subtle calculus of recognition by "context accumulation" worked its magic.

As the information shadows become thicker, more substantial, the need for explicit metadata diminishes. Our cameras, our microphones, are becoming the eyes and ears of the Web, our motion sensors, proximity sensors its proprioception, GPS its sense of location. Indeed, the baby is growing up. We are meeting the Internet, and it is us.

Sensors and monitoring programs are not acting alone, but in concert with their human partners. We teach our photo program to recognize faces that matter to us, we share news that we care about, we add tags to our tweets so that they can be grouped more easily. In adding value for ourselves, we are adding value to the social web as well. Our devices extend us, and we extend them.

Nor is this phenomenon limited to the consumer web. IBM's Smarter Planet initiative and the NASA-Cisco "planetary skin" project both show how deeply business will be transformed by the sensor web. Oil refineries, steel mills, factories, and supply chains are being instrumented with sensors and exactly the same kind of machine learning algorithms that we see in web applications.

But as is so often the case, the future isn't clearest in the pronouncements of big companies but in the clever optimizations of early adopters and "alpha geeks." Radar blogger Nat Torkington tells the story of a taxi driver he met in Wellington, NZ, who kept logs of six weeks of pickups (GPS, weather, passenger, and three other variables), fed them into his computer, and did some analysis to figure out where he should be at any given point in the day to maximize his take. As a result, he's making a very nice living with much less work than other taxi drivers. Instrumenting the world pays off.

Data analysis, visualization, and other techniques for seeing patterns in data are going to be an increasingly valuable skillset. Employers take notice.

This isn't to say that there isn't a huge role for unique identifiers for objects, especially fungible objects that are instances of a well-known class (like a book or music collection). But evidence shows that formal systems for adding a priori meaning to digital data are actually less powerful than informal systems that extract that meaning by feature recognition. An ISBN provides a unique identifier for a book, but a title + author gets you close enough.

Projects to systematically categorize raw sensor data may be created, along the lines of the Astrometry project, whose founders claim, "We are building an 'astrometry engine' to create correct, standards-compliant astrometric meta data for every useful astronomical image ever taken, past and future, in any state of archival disarray." Using this engine, the Flickr astrotagger bot trolls Flickr for images of astronomical objects and gives them proper metadata, which then allows them to be included in astronomical image search by name. This is a service directly analogous to CDDB: a lookup service that maps messy sensor data to a regularized lookup database.

As is often the case, the early examples are often the work of enthusiasts. But they herald a world in which entrepreneurs apply the same principles to new business opportunities. As more and more of our world is sensor-enabled, there will be surprising revelations in how much meaning—and value—can be extracted from their data streams.

Consider the so-called "smart electrical grid." Gavin Starks, the founder of AMEE, a neutral web-services back-end for energy-related sensor data, noted that researchers combing the smart meter data from 1.2 million homes in the UK have already discovered that each device in the home has a unique energy signature. It is possible to determine not only the wattage being drawn by the device, but the make and model of each major appliance within—think CDDB for appliances and consumer electronics!

Mapping from unstructured data to structured data sets will be a key Web Squared competency.

The Rise of Real Time: A Collective Mind

As it becomes more conversational, search has also gotten faster. Blogging added tens of millions of sites that needed to be crawled daily or even hourly, but microblogging requires instantaneous update—which means a significant shift in both infrastructure and approach. Anyone who searches Twitter on a trending topic has to be struck by the message: "See what's happening right now" followed, a few moments later by "42 more results since you started searching. Refresh to see them."

What's more, users are continuing to co-evolve with our search systems. Take hashtags on Twitter: a human convention that facilitates real-time search on shared events. Once again, you see how human participation adds a layer of structure—rough and inconsistent as it is—to the raw data stream.

Real-time search encourages real-time response. Retweeted "information cascades" spread breaking news across Twitter in moments, making it the earliest source for many people to learn about what's just happened. And again, this is just the beginning. With services like Twitter and Facebook's status updates, a new data source has been added to the Web—real-time indications of what is on our collective mind.

Guatemala and Iran have both recently felt the Twitter effect, as political protests have been kicked off and coordinated via Twitter.

Which leads us to a timely debate: There are many who worry about the dehumanizing effect of technology. We share that worry, but also see the countertrend, that communication binds us together, gives us shared context, and ultimately shared identity.

Twitter also teaches us something important about how applications adapt to devices. Tweets are limited to 140 characters; the very limits of Twitter have led to an outpouring of innovation. Twitter users developed shorthand (@username, #hashtag, $stockticker), which Twitter clients soon turned into clickable links. URL shorteners for traditional web links became popular, and soon realized that the database of clicked links enable new real-time analytics. Bit.ly, for example, shows the number of clicks your links generate in real time.

As a result, there's a new information layer being built around Twitter that could grow up to rival the services that have become so central to the Web: search, analytics, and social networks. Twitter also provides an object lesson to mobile providers about what can happen when you provide APIs. Lessons from the Twitter application ecosystem could show opportunities for SMS and other mobile services, or it could grow up to replace them.

Real time is not limited to social media or mobile. Much as Google realized that a link is a vote, WalMart realized that a customer purchasing an item is a vote, and the cash register is a sensor counting that vote. Real-time feedback loops drive inventory. WalMart may not be a Web 2.0 company, but they are without doubt a Web Squared company: one whose operations are so infused with IT, so innately driven by data from their customers, that it provides them immense competitive advantage. One of the great Web Squared opportunities is providing this kind of real-time intelligence to smaller retailers without monolithic supply chains.

As explained so eloquently by Vivek Ranadive, founder and CEO of Tibco, in Malcolm Gladwell's recent New Yorker profile:

> *"Everything in the world is now real time. So when a certain type of shoe isn't selling at your corner shop, it's not six months before the guy in China finds out. It's almost instantaneous, thanks to my software."*

Even without sensor-driven purchasing, real-time information is having a huge impact on business. When your customers are declaring their intent all over the Web (and on Twitter)—either through their actions or their words, companies must both listen and join the conversation. Comcast has changed its customer service approach using Twitter; other companies are following suit.

Another striking story we've recently heard about a real-time feedback loop is the Houdini system used by the Obama campaign to remove voters from the Get Out the Vote calling list as soon as they had actually voted. Poll watchers in key districts reported in as they saw names crossed off the voter lists; these were then made to "disappear" from the calling lists that were being provided to volunteers. (Hence the name Houdini.)

Houdini is Amazon's Mechanical Turk writ large: one group of volunteers acting as sensors, multiple real-time data queues being synchronized and used to affect the instructions for another group of volunteers being used as actuators in that same system.

Businesses must learn to harness real-time data as key signals that inform a far more efficient feedback loop for product development, customer service, and resource allocation.

In Conclusion: The Stuff That Matters

All of this is in many ways a preamble to what may be the most important part of the Web Squared opportunity.

The new direction for the Web, its collision course with the physical world, opens enormous new possibilities for business, and enormous new possibilities to make a difference on the world's most pressing problems.

There are already hundreds of examples of this happening (see Call for Examples below). But there are many other areas in which we need to see a lot more progress—from our energy ecosystem to our approach to healthcare. Not to mention our financial system, which is in disarray. Even in a pro-regulatory environment, the regulators in government are hopelessly outclassed by real-time automated financial systems. What have we learned from the consumer internet that could become the basis for a new 21st century financial regulatory system? We need machine learning to be applied here, algorithms to detect anomalies, transparency that allows auditing by anyone who cares, not just by overworked understaffed regulators.

When we started the Web 2.0 events, we stated that "the Web is a platform." Since then, thousands of businesses and millions of lives have been changed by the products and services built on that platform.

But 2009 marks a pivot point in the history of the Web. It's time to leverage the true power of the platform we've built. The Web is no longer an industry unto itself—the Web is now the world.

And the world needs our help.

If we are going to solve the world's most pressing problems, we must put the power of the Web to work—its technologies, its business models, and perhaps most importantly, its philosophies of openness, collective intelligence, and transparency. And to do that, we must take the Web to another level. We can't afford incremental evolution anymore.

It's time for the Web to engage the real world. Web meets World—that's Web Squared. ▪▪

A Call For Examples!

As part of this paper and our work on the agenda for the Web 2.0 Summit, we'd like your input. We're looking to create a list of applications, services, and projects that reflect the Web Squared theme. A few examples:

- The election of Barack Obama has demonstrated how the Internet can be used to transform politics. Now, his administration is committed to exploring how it might be used to transform the actual practice of governing.

  The US Federal government has made a major commitment to transparency and open data. Data.gov now hosts more than 100,000 data feeds from US government sources, and the White House blog is considering a commitment to the 8 Open Data Principles articulated by a group of open data activists in late 2007. There's a celebration of the successes that many are now calling "Government 2.0." We'd love to hear about Government 2.0 success stories from around the world.

  But in his advice on the direction of the Government 2.0 Summit, Federal CTO Aneesh Chopra has urged us not to focus on the successes of Web 2.0 in government, but rather on the unsolved problems. How can the technology community help with such problems as tracking the progress of the economic stimulus package in creating new jobs? How can it speed our progress towards energy independence and a reduction in CO_2 emissions? How can it help us remake our education system to produce a more competitive workforce? How can it help us reduce the ballooning costs of healthcare?

- Twitter is being used to report news of disasters, and to coordinate emergency response. Initiatives like InSTEDD (Innovative Support to Emergencies, Diseases, and Disasters) take this trend and amp it up. InSTEDD uses collective intelligence techniques to mine sources like SMS messages (e.g., Geochat), RSS feeds, email lists (e.g., ProMed, Veratect, HealthMap, Biocaster, EpiSpider), OpenROSA, Map Sync, Epi Info™, documents, web pages, electronic medical records (e.g., OpenMRS), animal disease data (e.g., OIE, AVRI hotline), environmental feed, (e.g., NASA remote sensing, etc.) for signals of emerging diseases.

  The Global Virus Forecasting Initiative (GVFI) now deliberately collects data (in this case, about emerging diseases crossing over from animal to human) that can be fed into this global early-warning system.

- Our health care system is tottering. Meanwhile, there is little correlation between spending and outcomes. As Atul Gawande wrote in the New Yorker:

  "Local executives for hospitals and clinics and home-health agencies understand their growth rate and their market share; they know whether they are losing money or making money. They know that if their doctors bring in enough business—surgery, imaging, home-nursing referrals—they make money; and if they get the doctors to bring in more, they make more. But they have only the vaguest notion of whether the doctors are making their communities as healthy as they can, or whether they are more or less efficient than their counterparts elsewhere."

  In short, we're measuring the wrong things. How do we apply the lessons of Web 2.0 to measure the right things about healthcare?

- Companies like PatientsLikeMe and 23andMe are applying crowdsourcing to build databases of use to the personalized medicine community. 23andMe provides genetic testing for personal use, but their long term goal is to provide a database of genetic information that members could voluntarily provide to researchers. PatientsLikeMe has created a social network for people with various life-changing diseases; by sharing details of treatment—what's working and what's not—they are in effect providing a basis for the world's largest longitudinal medical outcome testing service. What other creative applications of Web 2.0 technology are you seeing to advance the state of the art in healthcare?

- How do we create economic opportunities in reducing the cost of healthcare? As Stanford's Abraham Verghese writes, the reason it's so hard to cut healthcare costs is that "a dollar spent on medical care is a dollar of income for someone." We can't just cut costs. We need to find ways to make money by cutting costs. In this regard, we're intrigued by startups like CVsim, a cardiovascular simulation company. Increasingly accurate data from CAT scans, coupled with blood flow simulation software running on a cloud platform, makes it conceivable to improve health outcomes and reduce costs while shrinking a multi-billion dollar market for angiography, an expensive and risky medical procedure. If CVsim succeeds in this goal, they'll build a huge company while shrinking the nation's healthcare bill. What other similar opportunities are there for technology to replace older, less effective medical procedures with newer ones that are potentially more effective while costing less?

- As part of the financial stimulus package, the government is spending $5 billion on weatherization subsidies. How might Web 2.0 technologies tell us if the program is meeting its goal of creating jobs and reducing energy usage?

- Forward looking companies are adopting real-time monitoring and management to build smarter supply chains, manage remote resources, and in general, improve their return on investment using what Doug Standley at Deloitte calls "Asset Intelligence." We'd love to hear examples from people who are deploying these technologies.

- Real-time traffic monitoring systems like Microsoft Clearflow reduce wasted time and energy commuting. Web services reporting progress of buses and trains against their scheduled times make public transit more effective and enjoyable. These are tangible consumer benefits from instrumenting the world. Sensor-driven congestion pricing schemes like the one IBM built for the city of Stockholm create economic incentives to reduce traffic at peak times. These initiatives also raise privacy issues. We're interested in hearing about success stories—and scare stories—about the way that instrumenting the world changes the way we live.

- Smart Grid initiatives will reduce our energy usage by increasing the intelligence of the system used to deliver it. As hinted at above, though, they will also open a whole new front in the war on privacy. The data that will be revealed by smart grid applications will not only make our utilities smarter, it will likely make marketers a lot smarter too. It is unlikely, though, to make them more humane and less intrusive!

About the Authors:

Tim O'Reilly
Founder and CEO, O'Reilly Media

O'Reilly is an activist for open source, open standards, and sensible intellectual property laws. Since 1978, O'Reilly has led the company's pursuit of its core goal: to be a catalyst for technology change by capturing and transmitting the knowledge of "alpha geeks" and other innovators. His active engagement with technology communities drives both the company's product development and its marketing. O'Reilly has built a culture where advocacy, meme-making, and evangelism are key tenets of the business philosophy.

John Battelle
Founder/Chairman/CEO, Federated Media

Battelle is an entrepreneur, journalist, professor, and author who has founded or co-founded scores of online, conference, magazine, and other media businesses. Previously, Battelle occupied the Bloomberg chair in Business Journalism for the Graduate School of Journalism at the University of California, Berkeley. He was founder, Chairman, and CEO of Standard Media International (SMI), publisher of *The Industry Standard* and TheStandard.com. Battelle recently wrote *The Search: How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture* (Penguin/Portfolio). He created Searchblog, a daily site which covers the intersection of media, technology, and the Internet. He is an expert in the field of search and has appeared on many national and international news channels.

About Web 2.0 Summit

Since 2004, Web 2.0 Summit (formerly Web 2.0 Conference) has brought the intelligence, innovation, and leadership of the Internet industry together in one place at one time. Through incisive plenary sessions, cut-through-the-hype onstage conversations, rapid-fire "high order bits" and "show me" presentations, visionaries and executives across key industries present their unique perspective on the Web's future-in-flux and how the tools and principles of Web 2.0 are impacting their businesses. Web 2.0 Summit brings to light how to put the power of the Web to work—its technologies, its business models, and perhaps most importantly, its philosophies of openness, collective intelligence, and transparency.

Web 2.0 Summit is brought to you in partnership with O'Reilly Media, Inc. and TechWeb, and is moderated by John Battelle, Program Chair, and O'Reilly CEO and founder, Tim O'Reilly.

Attendance at Web 2.0 Summit is limited to maintain an intimate setting and foster dialogue among all participants. Registration is by invitation only.

For more information, visit web2summit.com