

Andreas Weigend: How the Social Data Revolution Changes (Almost) Everything
Speech given at the World Innovation Forum in New York on June 8, 2010.

WORLD INNOVATION FORUM

June 8-9, 2010 | Nokia Theatre | New York

How the Social Data Revolution Changes (Almost) Everything

June 8, 2010

Andreas S. Weigend, Ph.D.

info: weigend.com

twitter: [@aweigend](https://twitter.com/aweigend)

email: andreas@weigend.com

Audio of the talk, slides and transcript are at weigend.com/blog/archives/210

Alright we have a lot to cover and I want to start with who of you has checked in here with Foursquare, Godwalla, etc.? So in the next 15 minutes I will change the way you think about data. I will think about data; how they're created, and think about how data is the digital air we breathe. How will this data which we create be shared?

Everybody is a publisher. Every company is a publisher, and it's about building an ecosystem and participating, for all of you as individuals, and as companies in these ecosystems. Lastly, what is the intended use, and maybe what is the unintended use of those data?

The first part of my talk, I want to talk to you about language change, behavior change, and the language – how we move from collecting, soliciting, mining, and segmenting data and people; from sniffing the digital exhaust if you will, to sharing, to distributing, to interpreting data, and by doing so, to actually empowering employees as well as outside people, helping our customers to help us.

Let's take a step back in history. Think about some of the technologies that enabled innovation. In the 1800's, it was that we learned to transport energy. That gave rise to the Industrial Revolution, as we all know. In the 1900's we learned to transport bits, which gave rise to the Information Revolution. What I call the Social Data Revolution is what I think is happening now; the same thing that happened to the transportation of data is now happening to the creation of data. The economics has basically become free to create and share data.

andreas@weigend.com | [@aweigend](https://twitter.com/aweigend) | weigend.com/blog/archives/210 | weigend.com

Andreas Weigend: How the Social Data Revolution Changes (Almost) Everything
Speech given at the World Innovation Forum in New York on June 8, 2010.

In the last 60 seconds alone, the last minute, about 4 million web searches happened across the search engines. Half a million pieces of content were shared on Facebook. About 100 thousand product searches were done at Amazon.com, 40 thousand tweets were created, and 40 thousand links were shortened. I see on your faces that you don't know what I'm talking about here so I'll tell you a little story from last night.

When I called up the guy who does the website for the World Innovation Forum, I told him stuff he didn't know. It was not by breaking the computer, but by looking at data which this company Bit.ly shares for everybody to see. What Bit.ly does is a small insert. Bit.ly says shorten, share, and track your links; not only your links, but also your competitors' links.

For instance here, this long URL gets shorted to a very short one, bit.ly/WIF2010. When you look at that, it turns out that 5,614 people clicked on that link. Bit.ly allows you to look at any of these shortened URLs and get all the statistics, as I said, yours and your competitors'.

The first homework for you is look at one of the bit.ly links you find on the web, add a plus (+) behind the link, and try to understand what is going on there. What I understood here was that almost all those 5,614 hits were done in one day. I then checked out what happened that day and sure enough, that's the day they send out the email. If you have some hypotheses about how marketing could be happening by looking at the public data, you can find out.

Ways of innovation in the last 20 years were search, how best to find stuff; social, think about Facebook which helps us to share stuff; and the mobile helps us to create stuff, some examples about how the Social Data Revolution changes almost everything.

We started by connecting computers. Then Google came along and said let's look at the connection between pages. As you know, page rank enabled Google, to look at the information webmasters have, not about their own page, but about other peoples' pages and to have a more relevant search than pages that just looked at the content within a page. Connecting computers, connecting pages, what are you connecting now?

Connecting people – what's underlying all of that? It's data. When you have a process which is an exponential process, the typical way of characterizing that process is saying what is the doubling time of that process? How long does it take to double the amount of data that all of us, as individuals, create? It's pretty shocking. The data each of us creates doubles every 1.5 to 2 years. That means after 5 years it's a factor of 10; after 10 years it's a factor of 100; after 20 years it's a factor of 10,000.

Twenty years ago, this information was created about me. That is the cover sheet of the folder the secret police in East Germany, the Stasi was keeping about me. Just think about the amount of work that went into compiling that information; whereas now, many of us reveal on Facebook what the KGB wouldn't have gotten out of us under torture. [laughter] So since then, there has been computation which was dramatically edited.

Andreas Weigend: How the Social Data Revolution Changes (Almost) Everything
Speech given at the World Innovation Forum in New York on June 8, 2010.

Communication – the economics of communication basically determines that sending stuff has become free for people, but now the main cost is the cost of the recipient. What does this mean for new innovation, by showing irrelevant stuff to people, as well as sensing?

Here is an example of an Adobe Flash Player. More than a billion of those are connected worldwide. "Do you allow Amazon web services to access your camera and your microphone?" Of course yes, because as we heard in the introduction, by knowing whether I'm sitting in front of my computer by myself or with a friend, it can show me much more relevant ads. Do I look fresh, do I look tired? For the microphone, what is the quality of my voice? Should I see different ads if it's late in the evening? Is it bright, or dark, etcetera?

Metro Group, a company in Germany I did consulting for, for their Future Store, has every single item with an RFID tag. As I might be walking through that store, grabbing that cream cheese, putting it back and going for the low-fat cottage cheese even if it tastes lousy; they know that I'm worried about losing a couple of pounds. As I continue down the aisle, that light from the top is coming down with arrows to the dieting pills, and by the way; the prices are not fixed. The prices change on the displays as I walk by.

Then they have something called "path intelligence" which knows how I'm walking through the store, based on my mobile. Other companies have whole malls instrumented by understanding how I walk through the mall. Maybe HSM should instrument the World Innovation Forum by knowing who hangs out with whom. Or cities – Oakland is an example where crime spotting data are made public. I was told that Wednesday is a day when there are no crime data in that street where the prostitutes hang out, which means that's the evening when the officers who tend to patrol there seem to be doing something else. People are very good at adaptive behavior and learning those patterns.

Let's look at insurance and specifically let's look here at car insurance. Pay as you drive, if you commit to putting a GPS device in your car, which carefully logs where you're going, then your rate will be computed based on you actually drive. I live in San Francisco and my car is in my garage. Nobody drives it so I'm not paying anything right now.

But, let me pose to you that if I'm driving out from Stanford late at night, no cars around, and I do my 90 miles an hour, and if I had an accident the car insurance would say "Young man, you don't expect us to pay for your accident if you break the law." How would you feel about that?

Here is another example, Fitbit. Fitbit is a little device you clip on your belt and it measures how you walk, how you sleep – if you're sleeping with a belt, that is [laughter] and it is interesting how it influences behavior. Should I take a cab from the hotel to here? Well I'll actually walk because I get more steps logged onto my Fitbit.

More extremely, I now sleep almost every night with Zeo. In case you don't know Zeo, Zeo is a little device you clip on your head. The first night we didn't get along well because I'm not used to sleeping with wires. I couldn't fall asleep so it was blinking next to my bed saying "You're not sleeping well, you're not sleeping well" [laughter] and that didn't help. My record so far is 86 and every night I go to

Andreas Weigend: How the Social Data Revolution Changes (Almost) Everything
Speech given at the World Innovation Forum in New York on June 8, 2010.

bed I think "Maybe I should go to bed a little bit earlier and maybe I'll make it to 87 tomorrow." These are all examples about how instrumenting the world changes but changes in what timescale.

Think about the devices; let's think about your mobile phones right now. Who has a different mobile phone now from what they had a year ago? That is certainly more than half the audience. The time scale of data and technology is roughly one year. Social norms are about 10 years, and biology is maybe something like 100 thousand years.

C-to-B, "Talk to Chuck". If you go to my blog there is a recording with two grad students and myself. We saw these ads all over San Francisco that said "Talk to Chuck". So we called 1-800-forschwab and said "Hello, this is Andreas. I would like to talk to Chuck." [laughter] "Who is this?" "My name is Andreas Weigend." "What would you like to do?" "I'd like to talk to Chuck." "Who is this?" "This is ..." so you can listen to the recording. At the end he said "Oh, we are all Chuck." [laughter] That's not what we mean by C-to-B.

For a moment imagine that you know all the things people here have bought. You know all of their friends and you know all of their secret desires. What would you do? You want to help people make better decisions. Of course you all know Amazon Reviews and Amazon truly helps people make better decisions. It's all about the data. It's about consumers giving the data to Amazon, C-to-B, both explicit data like reviews, and implicit data, like purchases or clicks.

"People who bought X also bought Y." It's just one giant matrix of 100 million times 100 million entries, where whenever you buy two items within an hour, that cell corresponding to the intersection of those two items get incremented by one. What you see coming back is complements of things you might want to buy in addition because other people do that.

Similarly, Customers who viewed an item also viewed another item, but that doesn't give you complements. It gives you substitutes because it's early in the decision process. The best combination of those is helping to lubricate that purchasing funnel by showing people who are viewing that item what items other people eventually bought.

Social proof – put your money where your mouth is, and trust; who of you looks into the search logs your company actually creates? Your customers share with you by manually typing in what they're interested in. Another one of these small, cheap things to do in the next couple of days is to show me the last thousand things people looked for.

What have we talked about so far? We've talked about a number of data sources. These are data sources of intention, very powerful. That was the search boxes. We have data sources of attention; what do you pay attention to which ultimately is our scarcest good? We have data sources of situation which I didn't talk about very much, such as geolocation data, also the device you're carrying. And of course we have connection data, which brings me to the next point.

Not far away from here, a friend of mine who used to be at NYU did an experiment with a large phone company. He compared top statisticians who really know all about building segmentation models, with

Andreas Weigend: How the Social Data Revolution Changes (Almost) Everything
Speech given at the World Innovation Forum in New York on June 8, 2010.

something very simple; comparing it to who talks to him. The result is that the adoption rate of that new phone service went up by a factor of 5 by looking at connection data alone, compared to all the complicated stuff. That's a lesson for you.

In the '90s, people competed by algorithms, which wrenched the last bit out of those data, until they confessed. In the 2000's, the data has become cheap – sensing, communication, the cost is essentially zero. Now think about how you can get data that other people don't have. How can you get people to knowingly and willingly, C-to-B, share data with you? What you will see is that it's much cheaper to actually build simple models, such as sending something to somebody whose friend bought that phone product, is not as complicated.

Traditionally, businesses think about coming up with a message, optimizing the message, focus groups, etcetera, and then blasting it out to the consumers. The Super Bowl is an example. Now, you know that consumers talk. Amazon Review is one example.

The second part of my talk is now about C-to-C, consumers to consumers. What do people share with each other? I want to start with an Amazon example, Share The Love. Here is how it worked.

I buy a book and as I purchase that book, a screen comes up and says "Hey Andreas, do you have any friends who might be interested in learning about that book?" I think yeah "so give us the email addresses." Now I enter the email addresses and they each get an email address saying "Your pal Andreas just bought this book and thought you might be interested. If you want to buy it, click here and you will get a 10% discount. Andreas gets 10% credit."

I appear smart to them because after all, I just bought a book. [laughter] Why does this have amazing conversion rates? The four C's of marketing are all hit: content – that book; context – I actually just put my money down; connection – I'm determining who I am picking, and finally conversation – it provides a reason for that long-lost friend to say "Hey Andreas, I saw you bought that book."

Here is a question for you; what is the essence of Facebook? It's distribution, that "share" button. I started off as an EEE and they told me that the purpose of communication is to transmit information. But then, is information just an excuse for communication? What does this mean for innovation? Those who can distribute their stuff and show the world what they're doing are the ones who are actually going to innovate much more quickly than those who get locked up in their organizations.

Here is an example. There is a "like" button, not the "like" button of the old style, where something gets bubbled up. No, it's a Facebook "like" button. As I click this, he or she knows that I and 132 other people liked that item, so I distribute my liking of something throughout the web.

Another example is if you go to weigend.com/blog is Facebook Connect to comment on my blog. Why would you comment on my blog? Because it gets distributed to your friends. You don't just do it because you feel you want to tell me something; you do it because that way all your friends know that you have something interesting to say about whatever you read on weigend.com/blog.

Andreas Weigend: How the Social Data Revolution Changes (Almost) Everything
Speech given at the World Innovation Forum in New York on June 8, 2010.

We now have those implicit data; who is connected with whom, and there are two uses for pretty much all businesses which come out of it. One is to make more money. The other one is to lose less money, as simple as that.

Social graph targeting is an example where birds of a feather buy together; where if I like to stay at the Four Seasons, probably my friends also like to stay at the Four Seasons. By going to a company who analyzes the graph, they can provide prospects to you about who might be running at a high-end hotel chain. But birds of a feather don't only buy together. They also steal together.

Fraud reduction – I worked with an insurance company in the Midwest. They go to a company called Rapleaf and say "That fellow over there submitted that claim. How are his friends? Are they sort of shady characters or are they clean?" Rapleaf comes back and says "Don't worry about it." Less resources spent on investigating that claim. This was the part of C-to-C data some companies extract in order to give you the ability to act upon it in a very simple way.

Last part here, part three, is C-to-W, consumer to world. You all know Amazon.com is publicly sharing interests, namely wish lists. Here is an example of a wish list. These are books I'm interested in. Every now and then somebody might buy me a book or at least he shows what kind of person he is.

Nike+, people actually buy some device, put it in their shoe as they go running, and then share those runs with the world. How often would they do that? It turns out they do it three times per week.

When I talked earlier this year to DLD in Germany, to Trevor Edwards, he said "We're not in the business of keeping the media companies alive. We're in the business of connecting with consumers." I said "No, you're not. You're providing a platform so consumers can connect to each other." Ultimately I don't care that it happens to be a Nike, I care that I might meet other people who might want to go running with me, or whatever. It really is a consumer to world perspective, where they provide a platform.

The most popular platform of course for C-to-W is Twitter. In my Stanford class, the Social Data Revolution, we had a group which had a very interesting project trying to find out what time of day, what day of week, what wording etcetera are the rules which people should know to have the highest probability of getting retweeted. Why does Twitter work?

My belief is Twitter works because there is an illusion of an audience. I have 1,000 followers and I feel that if I say who I am having breakfast with, 1,000 people read it. But I did a few experiments and I'm deeply crushed. [laughter] I offered theater tickets and nobody clicked. I've offered concert tickets because I was playing and nobody clicked. I think Twitter works because there is an illusion of an audience that I think I have a lot of people following me, but ultimately I am not so sure. Again, it's very simple for you. Play with it. Try it out.

I did a workshop at a bank and one of the exercises we did there was that we said let's engage with one customer who mentions the name of that bank. We did a quick Twitter search. There was some person and she said "... sucks!" Within a couple of minutes, the head of consumer banking at this bank

Andreas Weigend: How the Social Data Revolution Changes (Almost) Everything
Speech given at the World Innovation Forum in New York on June 8, 2010.

managed to engage with her, and that girl was surprised. Suddenly you actually have people, at a point in time when they care, and you can reach them. It's very different from traditional marketing.

I did a project with BestBuy. BestBuy employs people who do nothing but listen to Twitter feeds, simple tools, not complicated, not expensive. You need a chair and a computer, and maybe a telephone. Then they see who is talking about BestBuy, or BestBuy products, or even more interesting, about competitors' products. Not everybody is happy about everything they buy. Then the BestBuy agent tweets to him or her and says "If you have a problem, let's help you. We have Geek Squad. If you want that computer problem fixed, don't worry about it, we can do this. By the way you get a 30% discount toward your next purchase at BestBuy." You go to where the people are, and people are in a C-to-W world, sharing with the world what they are interested in at that moment.

For the last part, I've put a couple of insights together before we move to the Q&A period. It's about instrumenting the world. It's about those who manage to bridge the digital and the physical.

If you think back on companies like the hotel I'm staying in, they think they have me held hostage as a customer. I show up after a long, delayed flight, half drunk, and they know my room number. They say "Oh we have this lovely room next to the elevator that still works on a Saturday evening at midnight in general." You don't own your customers anymore.

At least then you would say "We know more about our product than anybody else." If you make Nokia phones, I invite you to look up on Google who knows more about that phone you own, Nokia or the web. The bad news is you don't own the product.

The last hope was of course the "b" word, brand. "At least we own the brand," but you know the story by now; the brand is no longer owned by the company. They can try hard but ultimately we moved from controlled production for the masses to uncontrolled production by the masses.

What can you do? Here is a company called GetSatisfaction with the tagline "Customer service is the new marketing." You grab people when they actually want to contact you. In this case, somebody had a problem with a Nokia phone. In terms of relevance, you can push the button "I have that problem too" and it bubbles up the relevance.

This company takes previously private instances of you calling up somebody, and they try to find the answer in some manual, and then reading something back to you, and they make them public instances. If you have that problem, you invite the world to actually see that problem and to also potentially help. Can companies avoid it?

If there are problems on products, the world already knows, as any Google search for any problem will show you. What GetSatisfaction does for you is they actually surface it in order to put it in a knowledge base, layer upon layer. This makes it more powerful to have the outside world help you with your problems. Nokia took this seriously. Rand is now Nokia Beta Labs and that is a very interesting site where lots of people have lots of open idea they create and share.

Andreas Weigend: How the Social Data Revolution Changes (Almost) Everything
Speech given at the World Innovation Forum in New York on June 8, 2010.

Here is another example, Threadless, discussing consumer ideas for t-shirts. Another one, StackOverflow, for those of you more interested in computer sites, or flatseats.com.

This was a very smart move by the previously B-to-B company called SkyTracks. SkyTracks used to check out those flat beds by measuring how long they were, the angle and so forth, and then they realized there is a consumer web that steals their thunder. How did they react? They said "Be our guests. Put the comments on our site." Flatseats.com is by far the best information you can find anywhere about how that business class or first class seat on that aircraft actually is. It's consumer-generated content.

SeatGuru, probably most of you know by now, which tells you which seat have no reclining and stuff like this. Often SeatGuru knows more than the airline knows. For instance, have you ever had the case where you call up your favorite airline and they say "No worries, there is power," and then there is something taped over it. SeatGuru knows.

The most powerful example is TripKick. You look up a hotel, and it tells you what rooms to avoid. Of course, if you like a corner room, an oversized room, it tells you that as well. They know way more than Starwood's central reservations. [laughter]

A final twist is Groupon which is a new business model. They say "If you bring in at least 10 people, we'll give all of you a deal." It is based on communication, based on you actually getting a deal for yourself, and helping the company.

Group buying, getting a better deal, those roots are certainly deep in China where this has been going on for a long time. A friend of mine's wife bought a Toyota. So did about 14,000 people that day in Shanghai. I said "How is this possible?" To break down that cognitive load of what car to buy, they talk, they meet, they talk, and then they converge and say "That's the car we want to buy." Think about the buying power they have as they go to one Toyota dealer after another Toyota dealer in Shanghai and say "Can you make us a deal? We'll buy 14,000." [laughter]

To summarize I have two slides. We have moved from eBusiness, where traditionally the firm is in the center, web 1.0 if you will; to meBusiness where I'm in the center and I am reaching out to companies – sometimes we say this as moving from CRM, customer relationship management, to CMR, customer managed relationships; to weBusiness. It really is about community. It's about lots of examples I gave you, whether it's running together or buying together; that is really what I mean by moving from eBusiness, to meBusiness, to weBusiness.

To innovation, I want to make just two points clear. Don't worry about dead data. My friend who runs Lufthansa always talks about these data tombs they have. Analysis paralysis, people don't know what to do about them. It only gets worse the longer they wait.

To live data, live data come in a whole bunch of forms. For instance, live data come as knowingly and willingly contributed social data. But live data also come as data in response to an experiment. It's much easier that way. You agree on the metrics in the company, and once you learn the metrics anyone in the company can potentially run an experiment. The other one is don't just worry about your internal

Andreas Weigend: How the Social Data Revolution Changes (Almost) Everything
Speech given at the World Innovation Forum in New York on June 8, 2010.

data. Traditionally, people needed clearance to actually share stuff with the rest of the world. Just think about it the other way around. Think about it as you should have clearance if you don't share stuff with the rest of the world. Bill Joy, the former Chief Scientist of Sun once said "Most smart people don't work here."

I gave you a bunch of examples of new data sources where people knowingly and willingly collect data, share data, annotate data, and create META data. I walked you through the three categories of C-to-B, C-to-C, and C-to W opening up to the world. In each case, I gave you a couple of exercises you could do as you get back to the office later this week.

Time for questions, and by the way; if you want to know more about me, my Twitter ID is @aweigend, and my website is weigend.com. The Bit.ly which I created will have the answers from the survey and my slides.

Audience Question 1: Do you have any advice on how we can be authentic in the era of Social Data? What is the most important ingredient for a successful innovation strategy?

Andreas:Let's take the two questions in sequence. The first question was how to be authentic, and the answer is; be authentic. [laughter] I think people have a very good AF, authenticity filter, and people have very good ways of figuring it out.

Let me give you an example. From that same insurance company I mentioned before, they said "We are considering hiring that company to find out who are best agents are." I said "Let's look on Yelp." I entered the name of that insurance company, agent, San Francisco. There was Igor. Igor was rated 5 stars by dozens of people. The first person who rated Igor 5 stars was a woman who said she works at Apple and rated 7 restaurants, and even the same drycleaner I have. It's a lot of work for Igor to create all these reviews artificially. It makes no sense from an economic perspective.

What was the second question? [laughter]

The example I gave you of GetSatisfaction is basically free. There's always a freemium model. It's free and if you want to have additional analytics, or be able to mark company employees with a certain symbol, then you paid something like \$100 or \$200 a month. Cost is no longer the limitation. It's your creativity. It's you thinking about stuff. It's you being willing to imagine what you can do and then doing it.

Think about what people do with the data which they can create, which they can share. What apps can you give them? People are very curious about themselves. We don't really know where we stand, and by us being able to learn, comparing ourselves to others, that's actually what drives us to share more data. I do believe it's the apps which matter with the Social Data Revolution. It's about what the end user gets out of sharing those data.

Audience Question 2: Do you have any specific suggestion for traditional companies: how can we learn more about the culture change of the Social Data Revolution?

Andreas Weigend: How the Social Data Revolution Changes (Almost) Everything
Speech given at the World Innovation Forum in New York on June 8, 2010.

It actually is the DNA of the company, whether you have the DNA of the company which embraces creativity, trying out stuff, playing with yourself, playing with others, versus [laughter] – what's so funny? [laughter] Versus having a more top-down approach.

One thing BestBuy did was Blue Shirt Nation for empowering all the people to basically add whatever comments they wanted to add. Instead of having 100 people do content, we had 100,000 people doing content and then the next step was to empower the world, which was another factor of 1,000 to hundred million people doing content. It is not as complicated as you think. It really is doing the first step, playing with stuff, trying it out, thinking about what you want to do.

One thing I did with a Japanese company was reverse mentoring. Normally, younger people have an older mentor. Why do I teach at Stanford? Because I learn so much from my students. So if I transfer this to you, my last little piece of advice would be to find someone in your company who is 25. Sit down with him one hour a week or a month and just learn from them. See how they do the world. I'm always shocked at how much I can learn from young people.

Thank you very much. I wish you a great rest of the Innovation conference.

About the speaker:

Andreas Weigend studies people and the data they create. He works with companies that are eager to discover and tackle the pertinent questions and to develop strategies to realize the untapped power of data. His clients include Alibaba, Best Buy, Goldman Sachs, Lufthansa, Match.com, Nokia, Singtel, Thomson Reuters, the World Economic Forum, as well as exciting startups around the globe. Previously, as the Chief Scientist of Amazon.com, he helped build the customer-centric, measurement-focused culture that has become central to Amazon's success. Since 2003, he has been teaching The Social Data Revolution, Data Mining and E-Business at Stanford, and The Digital Networked Economy at Tsinghua in China.

Andreas shares his insights at top conferences, such as the World Business Forum 2009 in Milan. Known as a lively and engaging speaker, his main goal is to challenge the minds of his audience, helping them understand the irreversible impact the Social Data Revolution has on people, business, and society. He also gives speeches and workshops to the world's most innovative firms that combine cutting-edge ideas with his expertise on behavioral economics and vision for consumer-enabling technologies.

Andreas received his undergraduate education in Germany and Cambridge (UK), and his Ph.D. from Stanford University in physics. His career as a data scientist combined with his deep industry experience across information-intensive organizations allows him to successfully bridge the gap between academia and industry.

Andreas lives in San Francisco, Shanghai, and on weigend.com.
