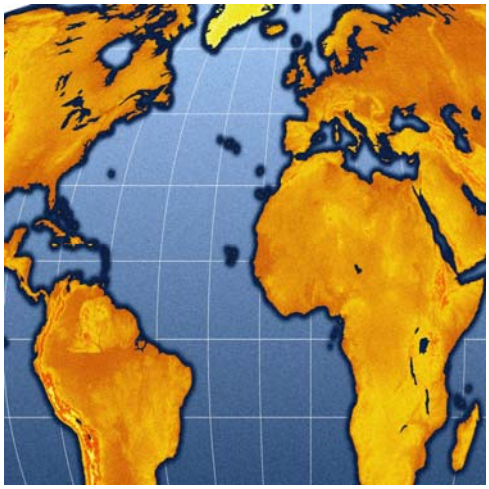
The Long Tail

An Interview with Chris Anderson

by Rita Koselka



A GBN WorldView Interview
June 2006

GBN Global Business Network
a member of the Monitor Group

The Long Tail: An Interview with Chris Anderson

Chris Anderson is the editor-in-chief of *Wired* magazine. Prior to taking over *Wired* in mid-2001, he was with *The Economist*. Besides a background in physics, Chris is a technology aficionado with a broad understanding of our changing economy and culture. *The Long Tail*, a book expanding his thesis from a celebrated December 2004 article on the "long tail" in economics will appear in July. In this interview, conducted by GBN's Rita Koselka, Chris explores the logic (and the math) behind his theory of infinite niche markets, explains why he thinks our human instinct for narrative is downright dangerous, and tells us why weak signals are the most powerful signals of all.

Rita Koselka: Your upcoming book *The Long Tail* is about a new paradigm of market dynamics. Historically, markets have generally featured fairly standard products that appeal to masses of people, which made their production and marketing economic and allowed producers to charge a reasonable price. But technology has made the production of many different items and the ability to reach audiences cheaper and easier, so that products that would have been uneconomic to produce or get to market not long ago are now making tons of money. You describe these new product markets as those found in the "long tail"—the portion of a statistical curve where the amplitude becomes very small. People have known about long-tail distributions for a long time, but you recognized that these distributions were relevant to product markets in a way that other people hadn't seen, and that defied the conventional wisdom. How did you realize that what you were seeing really represented a dramatic change? Did you have an "ah-ha" moment?

Chris Anderson: A long-tail distribution is not an unknown idea. It was intuitively grasped by Jeff Bezos, founder of Amazon, and everyone else who recognized what was different about the internet and who understood statistics. In science and economics, many functions have long tails, although they're often called fat-tailed or heavy-tailed distributions. One of the most common is what is often called a Pareto curve. It's the line on a graph that describes a situation in which a small number of things occur a lot and a large number of things occur a little, which is remarkably common. (Pareto first noticed it in wealth distribution, where 20 percent of Italians held 80 percent of the land, which led to the famous 80/20 rule). This is just basic statistics. But what Bezos and others did was to intuitively recognize that low-cost retail distribution allows you to go further down the curve, rather than just offering the most popular products.

For me, the "ah-ha" moment came when I was interviewing a guy named Robbie Vann-Adibé, who ran a digital jukebox company called eCast. The company makes jukeboxes with the usual neon and speakers, but rather than holding a hundred CDs or so it has a broadband connection and a hard drive.

He was asking me what percentage of the top 10,000 titles I thought sold at least once a month. First of all, that a jukebox could have 10,000 titles blew me away. But you might assume that the answer to his question would be 20 percent, because of the 80/20 rule. I assumed digital was different, so I went out on a limb and said 50 percent of the top 10,000 titles sell at least once a month. The answer turned out to be 98 percent. When you get something *that* wrong, you just sort of say, "Whoa." Everything they put out there sells at least a little. When I heard that, I knew

something here was different.

Then they went to 20,000 titles and 98 percent of them still sold. They went to 40,000 titles and 98 percent of them sold. I checked with Netflix and Amazon and in every instance the same was true: everything sells, sometimes just once a month, sometimes just twice. But if the marginal costs of manufacture and distribution are essentially zero, then all those ones and twos add up. When you graph it out you get a "power law" shape, just like a Pareto. As in traditional markets, you still have hits and non-hits. But what's different is that the sale of non-hit titles doesn't go to zero. And because the tail is so long, the area under that non-zero line, which is sometimes two orders of magnitude longer than the head, really adds up. It becomes 30 or 40 percent of the market. It's just the statistical recognition that when the tail is long enough, all those low-selling items can, in aggregate, amount to a big market—the economics of large numbers. This is counterintuitive because people don't really understand the economics of large numbers, but as we look closer at the evidence in the larger online businesses we can see it clearly showing up in the data.

Koselka: It reminds me of Wall Street trading or arbitrage. People don't appreciate how much money you can make on an eighth of a point.

Anderson: Exactly! Jeff Bezos started as a quant working on Wall Street, and quants—anyone with a math head—get intuitively what the integral of the function is. In calculus, an integral is the area under the curve. Bezos could intuit that the value or the volume under just the long-tail portion of the curve was large.

When you have an outlier piece of data like the one from eCast, one of two things is wrong. Either it's an anomaly—a weird artifact of digital jukeboxes—or you're looking at a significant effect that you just haven't focused on before. So I posed the same question, which no one had really asked, to Netflix and Rhapsody, another music service. What percentage of your titles actually sell? Then we reverse-engineered the question by looking at recent Amazon numbers. In each case, it worked out that 95 to 99 percent of their products sold at least once. Now *that* was interesting.

So we started summing up the numbers. We'd been looking at the left of the curve, because that's where the hits are. Now with infinite shelf-space markets, so you can have the hits *and* the niches. So we looked to the right, beyond the line where traditional retail ends. We started quantifying what we saw, just summing the curve. Wal-Mart ends at 4,500 CDs, Blockbuster ends at 3,000 DVDs, and Barnes & Noble ends at 100,000 books. So we asked, What is the aggregate size of the sales beyond that? And we came up with numbers like 30 percent and 35 percent. We said, "Whoa! There's a third of the market that's just dark matter that we hadn't measured before!"

Koselka: Do those sales have the same type of profit margins? Standard economics posits a law of diminishing returns.

Anderson: They are often higher margin sales, and there are two reasons for that. One is that the acquisition costs of niches are sometimes quite a lot less. Take DVDs, for example. You have a long tail of subject matter—the mainstream Blockbuster fare—then the niche fare and the older titles. The acquisition costs of brand-new DVDs produced this week will be about $19. An older one might be as little as $9. The price retailers can charge usually stabilizes around $15, so they tend to lose money in the first couple of weeks and make it up later. So, as you move down the tail, the margins get higher and higher, because the acquisition costs get lower but the price doesn't.

Likewise with books, Amazon discounts *The New York Times* bestseller list by 40 percent, but it doesn't discount the lower-selling. As long as the costs of providing these markets are near zero, you do very well; you basically have more pricing power with customers because it's harder for them to find these items.

Koselka: So let me try to get a larger view. Business has commonly focused on the hits, the big sellers. If we go from a hits-driven world to a non-hits-driven one, what does that mean? Will markets even out and mega-sellers become less common?

Anderson: You'll always have hits, but we're entering a world where they're no long the only products available. That doesn't mean that Pareto distributions, these power law curves of amplified inequality, never go away. They're an integral part of our world. They turn out to be common not just to economics and to markets, but biology is dominated by them. Earthquakes distribute in a power law. Cities distribute in a power law, particularly the smaller ones.

Koselka: Why?

Anderson: My theory on power laws, which comes loosely from an economics perspective, is that they are created by three components: variety of some type; inequality (they have differing value by some metric); and network effects (some feedback mechanism that amplifies whatever the value quality is). Word of mouth, for example, would be the feedback mechanism in markets like music. There are lots of different kinds of music. Depending on some discriminating quality, some music is better than others. Word of mouth amplifies the good and suppresses the bad, and this is why what would otherwise be just a straight line of sales becomes a power law.

Koselka: That's true for consumer products and products that are, roughly speaking, fashionable. What about other products where word of mouth may not be as important? Do sales of some products not correspond to power laws?

Anderson: All product markets that I know of show a power law distribution unless there is a factor that inhibits the market. Take Hollywood box office revenues, for example. What you discover is the scarcity effect in the carrying capacity of theatrical venues. There are a fixed number of movie screens, so where the data on box-office revenues hits zero is simply where they ran out of screens. Audiences still want to see the movie, but they just ran out of screens. So it's the scarcity effect. What this example tells you is that if you eliminate this bottleneck, there is an amount of latent market potential that's being suppressed by the scarcity function in the distribution. Go to any market and plot a logarithmic curve. If the market revenue distribution deviates from a straight line, there's a bottleneck. There are tremendous opportunities in eliminating bottlenecks.

There are other feedback mechanisms besides word of mouth that have network affects and account for why product markets obey power laws. A network effect describes how a product becomes more valuable to every individual as others use it; it's great if the people you hang out with listen to the same music you do. The classic example is the telephone. Once all the people you know have telephones that connect to one another, it is incredibly valuable to you.

The interesting question is: what is the network effect in nature? What do my three forces of power laws—variety, inequality, and network effects—mean in the context of earthquakes or cities or

species distribution? I think I can probably work it out for species distribution or why cities distribute the way they do, but what does it mean in the context of earthquakes?

Koselka: At least in economics, you could see data that corresponded to your model of a different market paradigm. Lots of areas of interest may not have clear data.

Anderson: Getting the data is always a challenge. Sarbanes-Oxley doesn't make it any easier, actually. The reality is that the economics of the twenty-first century are lying in the servers of Amazon and eBay and Google, which aren't always accessible. As with all rich data sets, you have to ask the right questions. It's noise if you ask the wrong question. But framing the question right and being able to figure out what would be the proper test of the model is hard. And sometimes the way to get new insight is to compare data sets that aren't really comparable. One thing we've been doing a lot is comparing offline and online markets, which is really, really tricky to do right.

Take music, for example. The offline market is a CD market, an album. The online market is a singles market. The data sets are very different. Online tend to be subscriptions, which is kind of an "all you can eat" model. The offline data is, obviously, based on a purchase or a la carte model. Books turn out to be tremendously difficult, not just because the book market is an archaic industry that isn't well tracked in any way, but because every statistic you see about the book industry is wrong. Amazon doesn't even know what's going on. We've been looking at Amazon's rankings, which we thought would be useful. They exhibit "hysteresis." Hysteresis is stickiness; things don't react to the forces on them proportionally. Amazon's algorithms stick. We're doing a long-tail time data series, and one album we've been tracking over time has been stuck at 721 for six weeks. It's just broken.

The Nielsen book-scan numbers don't track most of the market. None of the publishers actually know what's going on because they only see their own little part of it. The more we study these markets, the more we realize that we've got no real handle on the existing markets, to say nothing of the new ones. So there's an opportunity for the first time to do a real quantitative study of these things. Instead of looking from a book level or publisher level, we can look at the whole industry.

Koselka: You came to journalism and the business arena from an alternative background. You were a physicist. Jeff Bezos was a Wall Street quant, not a book seller. Is that significant?

Anderson: I think it is, actually. This is my next book, so it's not fully articulated yet, but the key success skill of the next era is going to be statistics. We're not wired to be able to really understand massively parallel statistical models. We're wired to understand linear narrative. We weight personal experience much higher than accumulated statistical data. My favorite quote—it's probably going to be the introductory quote to the book—is from Stalin, who said, "A single death is a tragedy. A million deaths is a statistic." The reality is that we weight personal experience and what happens around us so much more than the big numbers that we can't digest. This is why we're so drawn to the narrative. It's a linear stream, and our brains work linearly: beginning, middle, and end.

But the world is massively parallel. The reason that people missed the long tail is because they focused on the hits, which are a kind of compelling narrative: songs rushing up the charts. They paid no attention to the misses. This kind of hit-centric thinking and focusing on the winners—even thinking in terms of winners and losers—is such a blinkered way to see the world. Of course, you

do have winners and losers, but it's not like you can only have one. That's zero-sum thinking, or thinking the winners matter and the losers don't. It is this combination of zero-sum thinking, of hit-centric thinking, of linear thinking, that leads us to miss the weak signals.

It turns out that all the interesting stuff burbles up from the weak signals. Madonna released a new album and it is guaranteed to be a hit because it's going to have a lot marketing money behind it and lots of media coverage because Madonna is a big celebrity. That's not interesting. What's interesting is going to be what band is burbling up from MySpace, the online young adult site, right now.

So I think the key skill in the future is the ability to suppress the power of narrative and understand statistics in order to see the weak signals.

Koselka: There are people in different areas of business—from quantitative finance to business school professors to those pushing the use of probability tools for business planning—who are constantly pushing numeric models to businesses.

Anderson: It's not like you can just ask people to think statistically. We all have Excel, which has statistical functions, but most people just don't know what to ask. They don't even know how to understand the results. I think it's inevitable that most people won't ever get this, but a small number of people will. One of the keys to success is to find the people in your operation who do understand statistics, who can suppress narrative. Frankly, it's the geeks. In Asperger Syndrome, one of the symptoms is that sufferers have a hard time with empathy; they have a hard time understanding other people's stories. It's a gross generalization to say a lot of geeks have Asperger Syndrome, but I think you see the connection. In a sense, they're the canaries. They can sense something going on in the numbers and they're not blinded by the power of the hits and the power of the compelling narratives that everyone else is thinking about.

Narrative is very powerful—perhaps too powerful. The problem with narrative is that it sticks in your head, and you see the world through that narrative. But that narrative is just one story, and there are lots of different stories that don't all end the same. If you see the world through this one story, you tend to ignore the stories that don't fit into that model. And the problem is that the really important shift, the important surprise that comes to you, is not going to be the twist in your story; it's going to be a radically different story. You can't always imagine those radically different stories. What you have to do is kind of listen for them. And there are so many of them that you can't listen to them in a narrative sense; you have to listen to them statistically.

I'll give you an example. I think the most important advance in medicine, probably in all of history, is not anesthesia. It's not penicillin. It's the double-blind, controlled clinical trial, because before it we were driven by anecdotes and personal experience. We are so *compelled* by personal experiences and stories that this actually became medical practice. It wasn't until you could do something that suppressed narrative and privileged statistics that suddenly we knew what was real and wasn't real. So that was the introduction of statistics to medicine. And I think we now have to suppress our human instinct in a lot of other realms and let the number speak for themselves.

Koselka: As someone who manages a magazine and a bunch of journalists, how do you get them to think this way—to see the unseen story?

Anderson: I have such a poor answer, it'll be quick. All you can teach is rigor. You can teach basic numeracy, insist on looking for evidence, be skeptical, ask why or what if or what else could it be. It's just rigor, really. You take turns listening and then examining. It is not just fact-checking or asking around and hearing alternative things. It's actually in using your own powers of analysis to ask yourself: does this seem plausible? When I say "rigor," what I mean is thinking for yourself. It is allowing your understanding of the area you're writing about and your own ability to ask, at every point, is that true? How could I tell? Does that make sense? Does that add up?

Another thing that we do, which helps a lot in the rigor but is quite alien to American journalistic standards, is we don't attempt to be even-handed. We have a thesis. Thesis-driven journalism is something that's common around the world; it's certainly what I learned at *The Economist*. But it's alien in American journalism, which is driven by newspaper standards. Due again to the scarcity effects of information distribution back in their heyday, newspapers were the main source of information. So in American-style journalism, you basically have two sides to each argument and rather than use your analytical abilities, what you do is find someone who will say the opposite and then you offer the two with equal weighting.

At *Wired*, we don't do that at all. Instead, we have a thesis. It's more like the scientific method: generate a hypothesis and then design an experiment that you test. You collect data, and some of that data is numbers and some of that data is opinion and articulation and anecdote from people. But it's all basically testing the model. And, as in the case of an experiment, the model morphs. Your thesis evolves because you listen to the data. You're not single-minded and you're not agenda-driven. You're thesis-driven. There's a difference. The journalist is a sort of scientist and the sources represent the experiments—but this is a model of journalism that not everyone is comfortable with. They want *balance*. But stories aren't balanced. Typically, there's one that's right and one that's wrong and you shouldn't give them equal weight.

Koselka: Wait a minute. You were talking about the tyranny of the power of narrative but you laud the value of being thesis-driven. I think of a thesis as a framework for understanding cause and effect, which is basically what happens in narrative as well.

Anderson: I see a clear distinction between a narrative and a thesis. The idealized thesis is a mathematical model. Say you theorize that E equals MC squared. You'd then test that against the real world. The data would come back and it would either conform to the model or not, and that's how you test your thesis.

Science requires repeatability. So the narrative of a thesis is that you have to be able to do it again and come up with the same answer. There's obviously some similarity there. But I think the problem is that narrative without a strong model or a thesis underlying it takes on undue power. You expect that all stories will run the same way, and until you understand the underlying reasons why that story ran that way, you don't really know whether you have the confidence of prediction. Some stories are true and predictive and universal and some stories aren't. But because they're both compelling, we don't have a good way to tell one from the other.

Storytelling is a technique in economics, as well. There are equations, of course, but broadly you have a kind of an explanation of what happened and it involves a story. Economics has poor predictive power because the best stories tend to win, rather than the best data analysis. The data is complicated and the interpretation is ambiguous and so you end up with stories. Keynesian

economics is a good story. The invisible hand is a good story. Some of them are right; some of them are wrong. But the one that you subscribe to basically depends on which narrative compelled you.

The challenge in writing my book was to be able to go from an economic framework to a series of examples and anecdotes—dare I say narratives—that didn't do violence to the truth. I wrestle with this because every one of those stories I told in my article is wrong in a little way; every one of them is a caricature. The story's right, but it isn't a perfect example of the model. I feel obliged to qualify and to point out why this isn't representative, and that's not what people want to hear.

Malcolm Gladwell, the *New Yorker* writer and author of *The Tipping Point*, is the king of narrative-driven economics. He tells a narrative so brilliantly that no one sits back and asks, "Is that really generalizable? Is that the only way it can happen, or are there other ways?" It's just incredible power! Now it happens that he's right *and* tells brilliant narrative. But he tells such brilliant narrative that he wouldn't have to be right and it would still have equal impact. If you talk to people in consumer behavior, they feel *The Tipping Point* is exaggerated. I don't worry too much about that because I think that the book is broadly right and so powerfully told that it's a net positive.

Koselka: Let's think about the long tail in different terms than economic markets. One area that comes to mind is the market for ideas or ideologies. Is there a political, cultural, or intellectual analogy around a hits-driven world? And is that changing?

Anderson: In any instance where scarcity has limited the number of choices available and that scarcity is due to some structural effect that's changing, then yes. Whenever choice has been constrained by some bottleneck and that bottleneck is disappearing or it is widening, suddenly you have a massive increase in variety, in choice. More voices are heard or more ideas propagated to more people and to the market.

When you have a massive increase in variety, what you typically find is that it doesn't collapse back down to a hits thing. What you find is that almost all qualities in our world are distributed over a wide range, be they human qualities or intellectual qualities or economic qualities or cultural qualities. As scarcity functions disappear, that natural range of diversity emerges. You find that there are always going to be interesting things happening on the edges. As a matter of fact, *all* interesting things are first on the edges. And because you're taking these barriers to entry down, they have a way to propagate.

Koselka: But it doesn't lead to flattening, to a commonality of culture or political sentiment?

Anderson: It *never* leads to flattening. That's the amazing thing. I realize I sound very politically correct here, but diversity is not only the natural state of the world—it's also the optimum state of the world. Because we're a gregarious species and we're trained to think that we want to be like one another and live in lock step with our neighbors, diversity among humans seems strange. In some aspects of our lives we do act alike, but in everyone's life, in some part of their life, they're on the fringes. Everyone's got a seed of radical niche in them. And each one of our radical niches is different. When the bottlenecks disappear, you find that interesting things have the capacity to emerge, be heard, and then amplified.

Let me give you another example. You might imagine that a curve drawn around the distribution of

political sentiment of the American population today would look like a bell curve, right? Some people have fringe opinions but most people are nearer the center. Our political system of the moment, which is distorted by electoral inefficiencies, has reduced this to a camel curve of two humps—Democrats and Republicans. But you could have just as easily approximated that natural distribution with four peaks or 16 or whatever. How many parties in the right number of parties? What's the best way to approximate the full range of political diversity in the population?

Koselka: So among all this diversity on the fringes, in the tails, how would I identify what is important? How would you look into the tail to find out what's an interesting development?

Anderson: At this point, you fall back on information theory. It's about weak signals. The distortion of a two-party system dampens the weak signals of what large portions of the electorate want. The old assumption was that weak signals could be ignored because they can't propagate. The new assumption is that weak signals can't be ignored because they can so easily propagate. A phrase I use a lot in the book is that "ants have megaphones." If everyone has the power of propagation, you have to listen to the weak signals because the weak signals can turn into strong signals very, very quickly. And these new strong signals aren't just yesterday's signals that morphed; they can become something else very quickly.

I think the interesting thing is to get back to your point about what makes for power laws. In each domain, how do weak signals become strong signals? What are the means of amplification? How would you spot the snowball effect? What are the mechanisms by which they can snowball? They can be communications mechanisms, but also the human connections that allow or even encourage weak signals to become strong signals. How do musical bands move from the tail to the head of the curve?

To use a trivial entertainment example again, the old model of identifying musical talent was to have these guys called A&R guys—talent scouts. The listening ability of a record label was a function of how many talent scouts they had and how many nights there were in the year and how many bands they could see in a night.

Now you have the rise of bands that are burbling up from the grassroots. MySpace hits, such as "The Arctic Monkeys" and "Clap Your Hands, Say Yeah" and such. No talent scout discovered them; the mob discovered them. This is what I call "post-filtering" versus pre-filtering. Pre-filtering is the record label, who identifies the talent, guesses at what's going to be popular, and then *makes* it popular by putting the band in the studio and marketing the hell out of the album and putting it on shelves. Post-filtering is when everything's already out there. Rather than guess who's going to be popular, you *measure* who's popular.

What you're looking for is data. Sometimes that data is numbers and sometimes that data is news events and sometimes that data is what people say. What you do is compare that data with a model of what that data should be, and see if they agree. Data is noise unless you have a model to compare it with. This is the way the long tail started: I had a model of what I believed markets looked like. And then I heard a data point that was off, it was an outlier. Remember "The Matrix"? Neo sees the black cat twice, and says, "Whoa, deja vu?" That means that there's a glitch. It's a little thing, but it shouldn't happen. You have a model of what should happen and what shouldn't happen. Then something that shouldn't happen happens. Either something big is changing, which was the case in "The Matrix," or once we investigate we'll learn it really is trivial. The problem is you can waste a

lot of time following trivial outliers. But you need to start with a model of what the market should look like, listen for the outliers, and then have some way to evaluate them quickly.

Koselka: There's a tension, though, between having a theory or model that perhaps limits what data you look for and being open to listening for what comes up in the data having nothing to do with your theory.

Anderson: Looking for counterfactuals to your model is terribly important. But there is a special problem when the noise is so intense. Let me give you an analogy from science. In science, good theories need to be followed by experiment. Then, based on the data, you kill off the theories that didn't work and the theories that did work generate more theories. Then more experiments kill off some and so on. This is the natural progression of science.

Unfortunately, in some disciplines, the experiments have been unable to keep up because of the scarcity functions. In the case of physics, this is the current gap between the limits of our largest particle accelerators and the untestable predictions of string theory. When this gap grows, you have this really dangerous area. Experiment stops moving, theories have generated more theories that have generated more theories, and you get just total confusion. And in a chaotic world of too many competing theories with not enough evidence, guess what? The best "stories" tend to win.