

[Back to Blog](#)

How To Predict The Future

Jun 22, 2012

Category

The Future

#future

#avogadro

#predictions

#hertling

*Today's post is a guest post from William Hertling, author of the award-winning **Avogadro Corp: The Singularity Is Closer Than It Appears** and *A.I. Apocalypse*, near-term science-fiction novels about realistic ways strong AI might emerge. They've been called "frighteningly plausible", "tremendous", and "thought-provoking". By day he works on web and social media for HP. Follow him on twitter at **@hertling** or visit his blog **williamhertling.com**.*

*I'm a huge fan of William and his writing as you can see from **my review of his book Avogadro Corp**. So when William offered to write a guest post on how to predict the future, I enthusiastically said yes. Take a look – and take your time.*

Pretty much everyone would like a sure-fire way to predict the future. Maybe you're thinking about startups to invest in, or making decisions about where to place resources in your company. Maybe you just care about what things will be like in 10, 20, or 30 years.

There are many techniques to think logically about the future, to inspire idea creation, and to predict when future inventions will occur.

I'd like to share one technique that I've used successfully. It's proven accurate on many occasions. And it's the same technique that I've used, as a writer, to create realistic technothrillers

There seem to be two schools of thought on how to predict the future of information technology: looking at software or looking at hardware. I believe that looking at hardware curves is always simpler and more accurate.

This is the story of a spreadsheet I've been keeping for almost twenty years.

In the mid-1990s, a good friend of mine, Gene Kim (founder of Tripwire and author of **When IT Fails: A Business Novel**) and I were in graduate school together in the Computer Science program at the University of Arizona. A big technical challenge we studied was piping streaming video over networks. It was difficult because we had limited bandwidth to send the bits through, and limited processing power to compress and decompress the video. We needed improvements in video compression and in TCP/IP – the underlying protocol that essentially runs the Internet.

The funny thing was that no matter how many incremental improvements we made (there were dozens of people working on different angles of this), streaming video always seemed to be just around the corner. I heard "Next year will be the year for video" or similar refrains many times over the course of several years. Yet it never happened.

Around this time I started a spreadsheet, seeding it with all of the computers I'd owned over the years. I included their processing power, the size of their

I looked at the future predictions for “modem speed” (as I called it back then, today we’d called it internet connection speed or bandwidth). By this time, I was tired of hearing that streaming video was just around the corner, and I decided to forget about trying to predict advancements in software compression, and just look at the hardware trend. The hardware trend showed that internet connection speeds were increasing, and by 2005, the speed of the connection would be sufficient that we could reasonably stream video in real time without resorting to heroic amounts of video compression or miracles in internet protocols. Gene Kim laughed at my prediction.

Nine years later, in February 2005, YouTube arrived. Streaming video had finally made it.

The same spreadsheet also predicted we’d see a music downloading service in 1999 or 2000. Napster arrived in June, 1999.

The data has held surprisingly accurate over the long term. Using just two data points, the modem I had in 1986 and the modem I had in 1998, the spreadsheet predicts that I’d have a 25 megabit/second connection in 2012. As I currently have a 30 megabit/second connection, this is a very accurate 15 year prediction.

Why It Works Part One: Linear vs. Non-Linear

Without really understanding the concept, it turns out that what I was doing was using linear trends

breakthrough. This is what I mean by “forget about trying to predict advancements in software and just look at the hardware trend”.

It’s still necessary to imagine the future development (although the trends can help inspire ideas). What this technique does is let you map an idea to the underlying requirements to figure out when it will happen.

For example, it answers questions like these:

– *When will the last magnetic platter hard drive be manufactured?* 2016. I plotted the growth in capacity of magnetic platter hard drives and flash drives back in 2006 or so, and saw that flash would overtake magnetic media in 2016.

– *When will a general purpose computer be small enough to be implanted inside your brain?* 2030. Based on the continual shrinking of computers, by 2030 an entire computer will be the size of a pencil eraser, which would be easy to implant.

– *When will a general purpose computer be able to simulate human level intelligence?* Between 2024 and 2050, depending on which estimate of the complexity of human intelligence is selected, and the number of computers used to simulate it.

Wait, a second: Human level artificial intelligence by 2024? Gene Kim would laugh at this. Isn’t AI a really challenging field? Haven’t people been predicting artificial intelligence would be just around the corner for forty years?

SXSW, one of my co-panelists objected to the
How To Predict The Future - Brad Feld
notion that exponential growth in computer power was, by itself, all that was necessary to develop human level intelligence in computers. There are very difficult problems to solve in artificial intelligence, he said, and each of those problems requires effort by very talented researchers.

I don't disagree, but the world is a big place full of talented people. Open source and crowdsourcing principles are well understood: When you get enough talented people working on a problem, especially in an open way, progress comes quickly.

I wrote an article for the IEEE Spectrum called **The Future of Robotics and Artificial Intelligence is Open**. In it, I examine how the hobbyist community is now building inexpensive unmanned aerial vehicle auto-pilot hardware and software. What once cost \$20,000 and was produced by skilled researchers in a lab, now costs \$500 and is produced by hobbyists working part-time.

Once the hardware is capable enough, the invention is enabled. Before this point, it can't be done. You can't have a motor vehicle without a motor, for example.

As the capable hardware becomes widely available, the invention becomes inevitable, because it enters the realm of crowdsourcing: now hundreds or thousands of people can contribute to it. When enough people had enough bandwidth for sharing music, it was inevitable that someone,

IBM's Watson, which won Jeopardy, was built using three million dollars in hardware and had 2,880 processing cores. When that same amount of computer power is available in our personal computers (about 2025), we won't just have a team of researchers at IBM playing with advanced AI. We'll have hundreds of thousands of AI enthusiasts around the world contributing to an open source equivalent to Watson. Then AI will really take off.

(If you doubt that many people are interested, recall that more than 100,000 people registered for Stanford's free course on AI and a similar number registered for the machine learning / Google self-driving car class.)

Of course, this technique doesn't work for every class of innovation. Wikipedia was a tremendous invention in the process of knowledge curation, and it was dependent, in turn, on the invention of wikis. But it's hard to say, even with hindsight, that we could have predicted Wikipedia, let alone forecast when it would occur.

(If one had the idea of an crowd curated online knowledge system, you could apply the litmus test of internet connection rate to assess when there would be a viable number of contributors and users. A documentation system such as a wiki is useless without any way to access it. But I digress...)

Objection, Your Honor

1 μm — 1985
800 nm (0.8 μm) — 1989
600 nm (0.6 μm) — 1994
350 nm (0.35 μm) — 1995
250 nm (0.25 μm) — 1998
180 nm (0.18 μm) — 1999
130 nm (0.13 μm) — 2000
90 nm — 2002
65 nm — 2006
45 nm — 2008
32 nm — 2010
22 nm — 2011
16 nm — approx. 2013
11 nm — approx. 2015
6 nm — approx. 2020
4 nm — approx. 2022

Source: Wikipedia

A common objection is that linear trends won't continue to increase exponentially because we'll run into a fundamental limitation: e.g. for computer processing speeds, we'll run into the manufacturing limits for silicon, or the heat dissipation limit, or the signal propagation limit, etc.

I remember first reading statements like the above in the mid-1980s about the Intel 80386 processor. I think the statement was that they were using an 800 nm process for manufacturing the chips, but they were about to run into a fundamental limit and wouldn't be able to go much smaller. (Smaller equals faster in processor technology.)

But manufacturing technology has proceeded to get smaller and smaller. Limits are overcome, worked around, or solved by switching technology. For a long time, increases in processing power were due, in large part, to increases in clock speed. As that approach started to run into limits, we've added parallelism to achieve speed increases, using more processing cores and more execution threads per core. In the future, we may have graphene processors or quantum processors, but whatever the underlying

There are two ways I like to use this technique. The first is as a seed for brainstorming. By projecting out linear trends and having a solid understanding of where technology is going, it frees up creativity to generate ideas about what could happen with that technology.

It never occurred to me, for example, to think seriously about neural implant technology until I was looking at the physical size trend chart, and realized that neural implants would be feasible in the near future. And if they are technically feasible, then they are essentially inevitable.

What OS will they run? From what app store will I get my neural apps? Who will sell the advertising space in our brains? What else can we do with uber-powerful computers about the size of a penny?

The second way I like to use this technique is to check other people's assertions. There's a company called Lifenaut that is archiving data about people to provide a life-after-death personality simulation. It's a wonderfully compelling idea, but it's a little like video streaming in 1994: the hardware simply isn't there yet. If the earliest we're likely to see human-level AI is 2024, and even that would be on a cluster of 1,000+ computers, then it seems impossible that Lifenaut will be able to provide realistic personality simulation anytime before that.* On the other hand, if they have the commitment needed to keep

At a recent Science Fiction Science Fact panel, other panelists and most of the audience believed that strong AI was fifty years off, and brain augmentation technology was a hundred years away. That's so distant in time that the ideas then become things we don't need to think about. That seems a bit dangerous.

* The counter-argument frequently offered is "we'll implement it in software more efficiently than nature implements it in a brain." Sorry, but I'll bet on millions of years of evolution.

How To Do It

This article is *How To Predict The Future*, so now we've reached the how-to part. I'm going to show some spreadsheet calculations and formulas, but I promise they are fairly simple. There's three parts to the process: Calculate the annual increase in a technology trend, forecast the linear trend out, and then map future disruptions to the trend.

Step 1: Calculate the annual increase

It turns out that you can do this with just two data points, and it's pretty reliable. Here's an example using two personal computers, one from 1996 and one from 2011. You can see that cell B7 shows that computer processing power, in MIPS (millions of instructions per second), grew at a rate of 1.47x each year, over those 15 years.

6	Total Growth	328.52	=B3/B2
7	Rate of growth	1.47	=B4/(1/B5)

I like to use data related to technology I have, rather than technology that's limited to researchers in labs somewhere. Sure, there are supercomputers that are vastly more powerful than a personal computer, but I don't have those, and more importantly, they aren't open to crowdsourcing techniques.

I also like to calculate these figures myself, even though you can research similar data on the web. That's because the same basic principle can be applied to many different characteristics.

Step 2: Forecast the linear trend

The second step is to take the technology trend and predict it out over time. In this case we take the annual increase in advancement (B\$7 – previous screenshot), raised to an exponent of the number of elapsed years, and multiply it by the base level (B\$11). The formula displayed in cell C12 is the key one.

	A	B	C
10	Year	Expected MIPS	Formula
11	2011	177,730	=B3
12	2012	261,536	=B\$11*(B\$7^(A12-A\$11))
13	2013	384,860	
14	2014	566,335	
15	2015	833,382	
16	2020	5,750,410	
17	2025	39,678,324	
18	2030	273,783,840	
19	2035	1,889,131,989	
20	2040	13,035,172,840	
21	2050	620,620,015,637	

I also like to use a sanity check to ensure that what appears to be a trend really is one. The trick is to pick two data points in the past: one is as far back as you have good data for, the other is halfway to

1986 and a point in 1998 exactly predicts the
How To Predict The Future - Brad Feld
bandwidth I have in 2012. That's the ideal case.

Step 3: Mapping non-linear events to linear trend

The final step is to map disruptions to enabling technology. In the case of the streaming video example, I knew that a minimal quality video signal was composed of a resolution of 320 pixels wide by 200 pixels high by 16 frames per second with a minimum of 1 byte per pixel. I assumed an achievable amount for video compression: a compressed video signal would be 20% of the uncompressed size (a 5x reduction). The underlying requirement based on those assumptions was an available bandwidth of about 1.6mb/sec, which we would hit in 2005.

In the case of implantable computers, I assume that a computer of the size of a pencil eraser (1/4" cube) could easily be inserted into a human's skull. By looking at physical size of computers over time, we'll hit this by 2030:

	A	B	C
1	Year	Size (cubic inches)	Notes
2	1986	1782	Apple //e with two disk drives
3	2012	6.125	Motorola Droid 3
4			
5	Year diff	26	
6	Size diff	290.94	
7	Rate of shrinkage per year	1.24	
8			
9		Size	
10	2012	6.13	
11	2013	4.92	
12	2014	3.96	
13	2015	3.18	
14	2020	1.07	
15	2025	0.36	
16			less than 1/4 inch on a side cube. could easily fit in your skull.
	2030	0.12	
17	2035	0.04	
18	2040	0.01	

computers have tended to be big square boxes
How To Predict The Future - Brad Feld
constrained by the standardized form factor of components such as hard drives, optical drives, and power supplies. I chose to use computers I owned that were designed for compactness for their time. Also, I chose a 1996 Toshiba Portege 300CT for a sanity check: if I project the trend between the Apple //e and Portege forward, my Droid should be about 1 cubic inch, not 6. So this is not an ideal prediction to make, but it's still clues us in about the general direction and timing.

The predictions for human-level AI are more straightforward, but more difficult to display, because there's a range of assumptions for how difficult it will be to simulate human intelligence, and a range of projections depending on how many computers you can bring to bear on the problem. Combining three factors (time, brain complexity, available computers) doesn't make a nice 2-axis graph, but I have made the **full human-level AI spreadsheet available** to explore.

I'll leave you with a reminder of a few important caveats:

– *Not everything in life is subject to exponential improvements.*

– *Some trends, even those that appear to be consistent over time, will run into limits.* For example, it's clear that the rate of settling new land in the 1800s (a trend that was increasing over time) couldn't continue indefinitely since land is finite. But it's necessary to distinguish genuine hard limits (e.g. amount of land left to be settled) from the

– *Some trends run into negative feedback loops.* In the late 1890s, when all forms of personal and cargo transport depended on horses, there was a horse manure crisis. (Read **Gotham: The History of New York City to 1898.**) Had one plotted the trend over time, soon cities like New York were going to be buried under horse manure. Of course, that's a negative feedback loop: if the horse manure kept growing, at a certain point people would have left the city. As it turns out, the automobile solved the problem and enabled cities to keep growing.

So please keep in mind that this is a technique that works for a subset of technology, and it's always necessary to apply common sense. I've used it only for information technology predictions, but I'd be interested in hearing about other applications.

LOG IN WITH

OR SIGN UP WITH DISQUS ?

Name



Share

Best

Newest

Oldest

< Previous Post

Next Post >

Brad Feld



Enter your email here to subscribe to Feld Th

Subscribe

Legal

All rights reserved © 2004–2024 Feld Thoughts

