

by Marisa Henry

As an environmental engineering student, lover of economics, and detail oriented thinker (as certified by countless personality tests and a few BuzzFeed quizzes), I put a *lot* of faith in data. It is second nature for me to consider all the data at hand before making a decision as simple as what to eat for breakfast. Academically, my career goal (contributing to the designs of more sustainable water systems) will rely heavily on data from a wide range of disciplines.

But even with my extremely data-driven mind, I still get a little freaked out by data based systems like targeted advertisements. It's disconcerting that the carefully placed sneaker ads on my Facebook Newsfeed seem to taunt me for Googling "how to train for a marathon" with no intention of actually running unless there's a zombie apocalypse. I would love to say I resent the use of this information in advertisements, but these ads were made possible thanks to the same revolution that will help me reach my career goals: big data.

Big data is a term used to describe the large amount of digital data that is analyzed by smart software for trends and patterns. Big data analytics are creating a revolution in measurements and decision making in almost every field. Steve Lohr, author of the new book *Data-ism* and experienced technology reporter for The New York Times, addressed the balance between the risks and rewards of big data during a keynote lecture at the Dawn or Doom 2 conference hosted at Purdue University last week.

Lohr noted people have been collecting and using data for centuries, but the rate at which we collect data has increased exponentially and will continue to do so for the foreseeable future. The line between "data" and "big data" isn't universally defined and is a moving target as the quantities of data collected continue to double. At its most basic, big data characteristics include the "3 Vs": (1) volume (big data tracks continuous data streams and doesn't sample), (2) velocity (big data is usually tracked in real time), and (3) variety (big data pulls from many sources including text, audio, video, and sensors).

More important than the sheer volume of data being collected, big data is expanding the diversity of data available, finding a way to impact nearly every field. In healthcare, each beeping machine is tracking patient data and being used to develop methods of predicting patients at risk of serious health problems before more traditional symptoms arise. In finance, banks have used big data to look beyond credit scores and assess other loan qualifications quickly, increasing the number of loans given out. In my area of interest (water resources), big data is improving our ability to model and understand the complex water systems we've created. Satellite imagery and data track weather patterns and atmospheric conditions. Groundwater monitors alert officials of unsustainable aquifer depletion rates. Smart water meters are analyzed for insights into system inefficiencies.

Despite all of the positive applications of big data, there are certainly many challenges. Productive big data analytics starts by asking the right questions. Data should be used to inform

decisions, not handpicked to support opinions. This essentially boils down to being wary of the correlation versus causation framework.

Critics have also commented that data can't quantify everything. There is still validity to intuition and experience that isn't captured in digital data. Lohr cited Apple co-founder and former CEO Steve Jobs as an example. Jobs is famous for relying on his incredible instincts in leading Apple to success.

In addition, big data is quantifying the past and current states of data. While this is extremely useful for analysis, it brings up issues in using big data to predict the future events, or predictive analytics.

Finally, and perhaps most notably from a public perspective, big data pushes the boundaries of privacy issues. I am not the only consumer that feels uncertain about the targeted ads on my Facebook newsfeed. In 2012, the Pew Internet & American Life Project found that roughly two-thirds of internet users don't like targeted ads because they "do not like having their online behavior tracked and analyzed".

Despite the creepy feeling of uses like targeted ads, there will always be tradeoffs as society accommodates such revolutionary new technologies. As Lohr recounted, the first release of the commercial Kodak camera was perceived as a violation of privacy because anyone could take public photos easily. Today, nearly everyone has access to a smartphone capable of taking high quality images embedded with data like locations, and few people bat an eye.

So despite its drawbacks, big data is transforming the economics of discovery in nearly everything we do. There are challenges to this new technology, but society will adapt as it achieves a yin-yang balance between privacy, proper data use, and productive advances in measurements and decision making. There will almost certainly be a few moments of doom as society adjusts to the big data revolution, but the dawn always comes.