

Five Minutes with Professor Gary King: Transformational power of big data lies, pure and simple, in its analytics

LSE blogs.lse.ac.uk/impactofsocialsciences/2016/08/17/the-transformational-power-of-big-data-lies-pure-and-simple-in-its-analytics/

8/17/2016

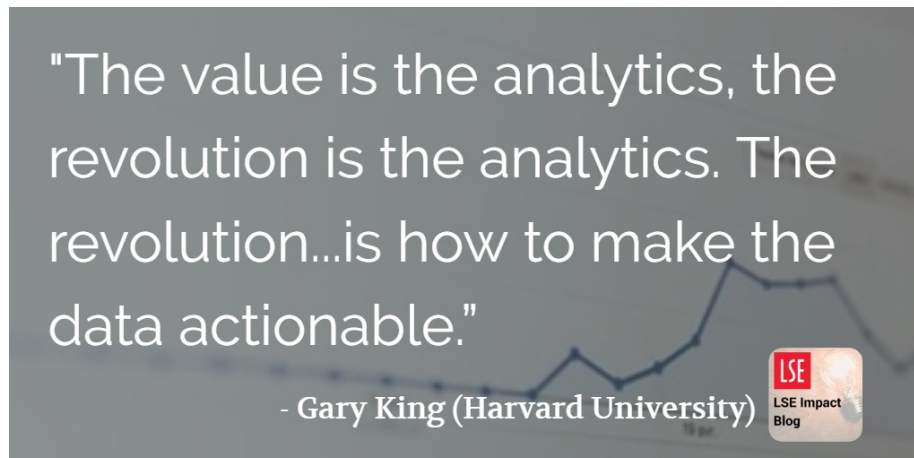
Michael Todd listened to a recent lecture by Gary King on the big data revolution in the social sciences. Professor King insists data is easy to come by and is in fact a by-product of the many improvements in information technology. The issue isn't its scale, volume or platform. It's what we make out of all of that and the analytical tools to do the job: "What's the big deal about big data...the answer is it's not about the data!"



The director of Harvard's Institute for Quantitative Social Science makes no bones about the utility of the term "big data." Gary King says the term helps the public "get" the revolution in commoditized data and the computational efforts involved in extracting value from that data. "My mom," he says, "now thinks she understands what I do." In a sense that should be intuitively obvious. And yet it's not.

King tells the story of a Harvard colleague who every year faced increasingly monstrous piles of data. One year the data exceeded what his computer could hold. The academic asked the university IT shop to "spec out a new computer," and the proposed bill for that cyber behemoth came back for \$2 million. King and a student "intercepted" this exchange and worked on crafting an algorithm "for almost two hours." The result? The initial colleague can now run his mountain of data on his laptop — and see results in 20 minutes or so.

"The most amazing thing about this story?" King prompts. "It's that it's not that amazing. It happens all the time. The innovation is the analytics." Even "off the shelf" analytics provide a huge improvement generating usable information compared to none, says King, but the astronomical leap comes from crafting custom analytical solutions — hardly a surprising statement from the head of a computational lab.



Data is easy to come by, he insists, and is in fact a by-product of improvements in information technology. Even if you choose to ignore this now commoditized flow, by the end of the year you'll still have more than you started the year with.

"What are you going to do with all that data? It's not that helpful, by itself, because you have to manage it. It's valuable, so you have to keep it. ... The value is the analytics, the revolution is the analytics. The revolution, that thing that we did not know how to do before, but that we are learning how to do now, is how to make the data actionable."

He cites Moore's Law, which predicts (successfully so far) that computer speed and power will double every 18 months. "That's nothing," King enthuses, compared to a competent grad student beaver away for an afternoon, who can create a thousand-fold increase by crafting algorithms to plow through these avalanches of data. That's why he's been an apostle for years of [restructuring the social sciences](#) so that can routinely accept and include "larger scale, collaborative, interdisciplinary, lab-style research teams."

And despite being a preacher, he's not a zealot when it comes to research methods. As he wrote in that same paper on restructuring social science:

"Fortunately, social scientists from both traditions are working together more often than ever before, because many of the new data sources meaningfully represent the focus and interests of both groups. The information collected by qualitative researchers, in the form of large quantities of field notes, video, audio, unstructured text, and many other sources, is now being recognized as valuable and actionable data sources for which new quantitative approaches are being developed and can be applied. At the same time, quantitative researchers are realizing that their approaches can be viewed or adapted to assist, rather than replace, the deep knowledge of qualitative researchers, and they are taking up the challenge of adding value to these additional richer data types."

He could, of course, use a little help in his crusade. At a recent event sponsored by SAGE Publishing in Washington, DC titled "The big deal about big data," King called on the policymakers and government officials in the audience to consider enacting a "treaty" on the collection, retention and sharing of big data that could serve the needs of government, academe and business while protecting the interests of the public.

During that event, which is captured in the video below, the academic offered policymakers real-world examples drawn from his own work about the value of signing on to this treaty. One of his highest-profile current examples focuses on deconstructing government reactions to social media in China. While [his most current work](#) shows how the Communist Party spoofs viral outbursts of social media activity to presumably influence public perceptions, at the Big Deal event he focused on earlier work showing how ordinary Chinese citizens work around government cyber roadblocks.

He explained that by comparing the pre-censored social media in China with post-censored social media, his team could “reverse engineer what the intentions of the censors were.” While the most engaging portion of his anecdote demonstrated how cyber-literate Chinese learn to speak their minds by coming up with new types of language to express forbidden sentiments, the policy-pertinent portion demonstrated that the Chinese government was less interested in shutting down grumbling and actually focused on preventing any forms of collective action.

His example neatly sums up King’s larger point. The issue isn’t the social media, its scale, volume or platform. It’s what we make out of all of that – and that requires the analytical tools to do the job. “What’s the big deal about big data?” he asks. “And the answer is it’s not about the data!”



Featured image: [Pixabay \(public domain CC0\)](#)

Note: This article gives the views of the author, and not the position of the LSE Impact blog, nor of the London School of Economics. Please review our [Comments Policy](#) if you have any concerns on posting a comment below.

About the Author

Michael Todd is the social science communications manager for SAGE Publishing.

Gary M. King is an American political scientist and quantitative methodologist. He is the Albert J. Weatherhead III University Professor and Director for the Institute for Quantitative Social Science at Harvard University.

- Copyright © The Author (or The Authors) - Unless otherwise stated, this work is licensed under a Creative Commons Attribution Unported 3.0 License.