

PROJECT SYNDICATE

THE WORLD'S OPINION PAGE

GLOBAL HEALTH & DEVELOPMENT



ALEX PENTLAND

Alex Pentland is Professor of Media Arts and Sciences, and Director of the Media Lab Entrepreneurship Program, at MIT Media Lab. He is the author of *Social Physics: Lessons from a New Science*.

MAY 27, 2014

Living Big Data

CAMBRIDGE – Big data is made from the digital trail that we leave behind when we use credit cards, mobile phones, or the Web. Used carefully and accurately, these data give us unprecedented scope to understand our society, and improve the way we live and work. But what works in theory may not translate well in the real world, where complex human interactions cannot always be captured, even by the most sophisticated models. Big data requires us to experiment on a big scale.

My own laboratory, for example, is building a Web site which, based on Google maps, uses society's digital trail to map poverty, infant mortality, crime rates, changes in GDP, and other social indicators, neighborhood by neighborhood – all of which will be updated daily. This new capability allows viewers to see, for example, where government initiatives are working or failing.

But, while such impressive visualization tools can dramatically enhance transparency and public knowledge, they are surprisingly limited when applied to solving society's problems. One reason is that such rich streams of data encourage spurious correlations.

Even the use of the normal scientific method no longer works; given so many measurements, and so many more potential connections among what's being measured, our standard statistical tools generate nonsensical results. Without knowing all possible alternatives, we cannot form a limited, testable set of clear hypotheses. And if we can no longer rely on laboratory experiments to test causality, we must test it in the real world, using massive volumes of real-time data. This involves moving beyond the closed, question-and-answer process typical of the lab, and applying our ideas in society, earlier and more frequently than ever before.

To see how things work in reality, we must construct living laboratories – that is, communities willing to try new ways of doing things (to be blunt, to act as guinea pigs). An example of such a living lab is the “open data city,” which I launched with the city of Trento in Italy, along with Telecom Italia, Telefónica, the research university Fondazione Bruno Kessler, the Institute for Data Driven Design, and local companies. Importantly, this living lab has the approval and informed consent of all involved; they understand that they are participating in a gigantic experiment whose goal is to create a better way of living.

One major challenge for a living lab is to protect individual privacy without diminishing the potential for better government. The Trento lab, for example, will pilot my proposed “New Deal on Data,” which gives users greater control over their personal data through trust-network software such as our open PDS (Personal Data Store) system. We hope that the ability to share data safely, while protecting privacy, will encourage individuals, companies, and governments to communicate their ideas widely, and so increase productivity and creativity across the entire city.

But the biggest difficulty in using big data to build a better society is being able to develop a human-scale, intuitive understanding of *social physics*. Although dense, continuous data and modern computation allow us to map many details about society, and to explain how communities might work, such raw mathematical models contain too many variables and complex relationships for most people to understand.

What is needed is some kind of dialogue between human intuition and the compelling reality of big data – a dialogue that is currently absent from management and government systems. If big data is to be deployed effectively, people must be able to understand and interpret the relevant statistics.

This calls for a new understanding of human behavior and social dynamics that goes beyond traditional economic and political models. Only by developing the science and language of social physics will we be able to make a world of big data a world in which we want to live.

<http://www.project-syndicate.org/commentary/alex-pentland-imagines-communities-guided-by-rich-streams-of-real-time-information>