

DATA AND CONTROL – A DIGITAL MANIFESTO

Wolfgang Pietsch, Munich Center for Technology in Society, TU München, Germany; pietsch@cvl-a.tum.de¹

This is a plea for democratic supervision and regulation of the large data sets that are currently being collected all over the digital world—a plea driven not by fears for the privacy of the individual, but by worries that a privileged knowledge of the mechanics governing the social world could allow for a one-sided and largely unrecognized control of the masses.

In the 19th century, in an age of unprecedented and seemingly limitless technological and scientific progress, scientists had a dream—a vision of a social science as systematic and powerful as the most developed of the natural sciences, a science of man modeled after physics. Like celestial mechanics, such a social science could be reduced to a few principles from which, under sufficient knowledge of the boundary conditions, everything else becomes derivable. It would allow for the prediction of the future course of the major actors in society just as celestial mechanics allows for determining the trajectories of Venus or Mars. The dream never materialized.

With obvious frustration, methodologists of the 19th and the 20th century, great and penetrating minds like the philosopher and economist John Stuart Mill, blamed the complexity of the social world and the fact that social phenomena do not fit neatly into a laboratory and are thus exempt from systematic experimentation. The social scientist cannot just exchange a democracy for a dictatorship, run history again, and record the effects of the change. Observing the fate of man in society rarely leads to causal knowledge allowing for prediction and control over social phenomena. Data collected in the social world can only disclose correlations, of which no one can be sure if and where they will hold again, not causal knowledge geared towards prediction and control, or so it seemed.

Then rose the digital age. In a world of big data and increasingly detailed recordings of the boundary conditions under which social phenomena happen, the known islands of causal knowledge are becoming larger, the sea of ignorance retracts. Almost certainly, these islands will become continents one not so distant day. This is not to claim that everything in the social sphere will become predictable, but prediction and control will be possible in ever more areas. After all, we are not dealing with a problem of principle, it is just a matter of quantity of data and of processing power. Complex phenomena require extensive experimentation, incredibly complex problems require incredibly large data sets. For a scientist in the 19th century, it was impossible to imagine how mankind could ever process or even possess such data sets.

It was beyond the imagination of the social scientist in the 19th century that one day the world would be populated by little machines equipped with various kinds of sensors and instruments, autonomously collecting information to transfer it to central archives, where the data would be analyzed and scanned for reliable relations. Your computers and laptops are such machines. And as the digital age advances further, the world will become ever more populated with sensing objects, phones, cars, screens, lampposts, that record an increasing variety of information. As the network of these parameters grows tighter around our daily lives, we are becoming, of necessity, increasingly predictable and controllable by those in possession of these data sets. Crucially, the data need not be personalized to your name, it suffices to know the relevant classes, to which a user belongs. In this respect, most traditional privacy debates go wrong.

The social physics of the digital age turns out a creature much different from the one imagined by our scientific forebears. It is not at all a celestial mechanics, not a neat and tidy construct, a

¹ Written in August 2012, forthcoming in *Public Culture*. This essay grew from work conducted within the Poiesis project. I acknowledge helpful discussions with my fellow members of the Milgram group, Nerea Calvillo, Orit Halpern, and Jesse LeCavalier.

hierarchical structure with a small number of fundamental principles. The social mechanics that is being recorded in the bodies of humming server farms, is not even a coherent, connected structure, it is a pluralistic entity, a patchy cluster of loosely related laws, arbitrarily overlapping, sometimes contradictory or even fully disconnected. More often than not, these laws hold only locally within a highly restricted range of boundary conditions, universal laws of sufficient generality are almost entirely lacking. It is thus not a scientific theory that fits into textbooks, it is an overwhelming collection of phenomenological data sets that can only live inside the large data storages of the computer era. It is a scientific theory that is never to leave its digital cage. But even though this creature was not what most scientists had expected, the theory is there and growing from day to day. And as it is getting more powerful we need to deal with it. The beast must be tamed.

Two kinds of control, two systems of laws govern the social world. One of them is well-documented in the codes of law that were built up over centuries and millennia of human coexistence in complex social and political environments. These normative laws are largely fixed by convention delimiting the freedom of the individual in her/his social interactions. They can be consciously broken by accepting the consequences laid down in law books—fines, prison, or even death. Our modern democratic societies have developed a highly complex and well-regulated process how these normative laws can be implemented and changed.

The other kind of laws is much less visible, they are veiled behind the illusion of unlimited human freedom and a mental reconstruction of the human condition, according to which all our actions are due to rational and conscious decisions. Psychologists have long pointed out the fallaciousness of this view, but some pre-programming of our minds prevents us from internalizing this fact. This other kind of laws much resembles the causal relations known from the natural sciences. Until the digital age, no considerable body of these laws was known, only some rather isolated social and psychological facts, exploited mainly by politicians and advertising companies. These laws were often weak and unreliable in little need of regulation. With the digital age, the situation is changing, the body of known reliable causal laws constantly growing. Unlike the normative laws, which largely concern situations in which we decide consciously, the causal laws govern the unconscious side of our thoughts and actions.

A colleague of mine recently complained about baby strollers popping up on his computer screen distracting him from the work he was supposed to be doing. Of course, he has a young child and they were recently on the lookout for strollers in the internet. Almost certainly, these are just the clumsy beginnings, some indication how the direction of our thoughts and steps may be controlled in a future world by placing triggers that increase our probabilities to think and behave in certain ways. And as the boundary between the digital and the physical dissolves, so does the boundary between advertising and unbranded reality. We are facing nothing short of the total commercialization of the social world. Gone are the times when publicity was forced into frames from which you could look away or which you could even switch off, when publicity was spatio-temporally constrained. Large internet companies use their causal knowledge to more or less unrecognizably steer our ways through the internet. The money that currently goes into the internet advertising business is proof enough that the principles work. What will happen when this knowledge is used for other things beyond making money? What damages are already inflicted on society by the commercialization of ever more areas of our social lives including the intrusion of advertising into our working environments? What will happen, when the internet and its principles will extend to the object-world, when the internet of things will be realized?

There are no easy answers, all wholesale-solutions are inappropriate. Open data leads to uncontrollable anarchy, while the current laissez-faire attitude, leaving internet firms to handle data at their will, leads to a knowledge oligarchy dangerously dominated by an arbitrary lot of a few large companies. We need a debate.