

There is gold in big data, but there are not enough gold miners

March 21 2018

The promise of big data has been on the horizon of the field of social sciences for years, but so far nobody has been able to deliver on this promise. In his inaugural lecture on 22 March, Professor and research methodologist Bernard Veldkamp explains why and offers solutions. His main point: we have to find different ways of dealing with this type of data. "The why-question should be replaced by pattern interpretation."

There is currently a revolution underway in terms of the amount of [data](#). Due to the rise of the internet, social media, mobile phones and all kinds of sensors, the amount of data rises by as much as forty percent annually. For [social scientists](#), who deal with the behaviour of individuals and groups, these enormous amounts of data can be a goldmine. However, even though the promise of big data has been on the horizon for the field of social sciences for many years, the analysis has not yielded the results anyone expected. Consequently, enthusiasm is slowly turning to scepticism. In the meantime, computer scientists who develop methods to analyse the data are making off with the goods. An important disadvantage, however, is that computer scientists, given their background, are unable to sufficiently take account of the context and the information they are studying. As a result, the usefulness of their findings is limited.

Closing the gap

According to Professor Veldkamp, it is time for a paradigm shift. As

one of the main reasons why social scientists lost the first battle, he mentions the fact that they tried to process and analyse the data using traditional methods for far too long. For traditional methods, the amount of data is simply too large. Moreover—and perhaps more importantly—the type of data involved is completely different compared to the data that was previously available. Because the data—unlike data from traditional samples, observations or questionnaires—has not been specifically collected for scientific analysis, the origin and quality are not always clear. Or, as Veldkamp simply puts it: "There is a lot of [white noise](#) in it." This is why he believes it is very important to choose a different approach. The field of research methodology, Veldkamp's field, is ideally suited for closing the gap between [big data](#) and the [field](#) of social sciences. "The data is worth its proverbial weight in gold, but there are currently not enough qualified gold miners for the job."

Patterns

Professor Veldkamp believes that the fact that the origin of the data is sometimes unknown and that there may be a lot of white noise does not necessarily have to be a problem, but it is something you need to factor into your analysis. Big data, precisely due to the large quantity, offers the possibility to statistically correct uncertainties. In addition, Professor Veldkamp believes it's essential to "model in the white noise". This means that you have to be more careful with assumptions and that any conclusions you draw must be less firm. Professor Veldkamp: "For this reason, it is becoming more important to interpret connections than to look at causality. Or, to put it more simply, pattern interpretation should replace the why-question."

Provided by University of Twente

Citation: There is gold in big data, but there are not enough gold miners (2018, March 21)
retrieved 9 April 2024 from <https://phys.org/news/2018-03-gold-big-miners.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.