

Event Review: Big Data – Small Devices

08.04.2016

<http://www.germaninnovation.org/news-and-events/past-events/past-event?id=f489a234-f89d-e511-bdec-00155dcfd969>

On March 7, 2016, the German Center for Research and Innovation (GCRI) New York and the University Alliance Ruhr (UA Ruhr) hosted a panel discussion on Big Data and resource-restricted systems. The speakers discussed the proliferation of digital information and its effect on society.

Prof. Dr. Katharina Morik, Head of the Collaborative Research Center SFB 876 and Professor of Computer Science at TU Dortmund University, kicked off the evening by presenting some insights from related research coming out of Germany. The basic research she is conducting not only focuses on improving data algorithms, but also on reducing devices' energy consumption. For example, she is currently working on smart containers for the logistics industry to make transportation and routing more efficient. Morik also spoke about smartphones, calling them a "human sensor," which detects human behavior by collecting the roughly 60 GB of data per year that is produced by each user. This results in a truly vast amount of data when one thinks of the billions of smartphone owners worldwide. Morik then presented the largest existing academic smartphone data set, which was created from 300,000 participants to provide insight into usage in terms of gender and app use behavior. Her main takeaway here was to address the issue of smartphone energy management. By predicting app usage patterns, her researchers are able to predict energy needs and thereby reduce energy expenditure. They aim to run their machine learning algorithms directly on smartphones to enhance this process as well as to ensure privacy. As algorithms themselves demand computing power, more energy-efficient algorithms will be needed in the future. Big Data analytics will make smartphones smarter and will even be run on restricted devices.

The next speaker, Prof. Dr. Kristian Kersting, Professor of Computer Science at TU Dortmund, spoke about the use of Big Data to analyze events, such as traffic jams. He presented the idea of combining physics, computer science, and communication networks through the use of Big Data to create models of conversational flow as well as simulations on individual driving behavior. The technology to gather this data may consist of traffic cameras, loop sensors built into the road, sensors built into cars, and cell phones acting as human sensors, which all feed these models with data. Together with machine learning and distribution over counts in a multivariate manner, Big Data is able to improve these physics models. Kersting called this a deep count learning model. He then provided a second example of the use of Big Data to highlight the phenomenon of viral videos and online firestorms that can cause damage to an organization's reputation. He showed the audience how to apply mathematical models to such events and to recognize their different stages. Issues like data jams caused by billions of active smartphone users worldwide pose important questions to society going forward. Kersting's takeaway was that, "almost everything that counts can be counted and almost everything that can be counted, counts."

The fourth speaker, Dr. Claudia Perlich, works for the New York-based start-up Distillery, which specializes in digital advertising. She spoke about the online advertising industry, which aims to connect digital content providers with business clients to show personalized ads to online users. It does so by tracking user behavior and gathering relevant data. Ads are sold in real time through ad exchange platforms that work like auctioneers – with roughly 50 billion bid requests each day at just her company alone. Perlich's job is to run digital ad campaigns for corporate clients. One question she grapples with on a daily basis is how exactly one finds the right customers to show an ad.

Companies want to see results for the money they spend on digital advertising. Their main goal is to create “conversions,” which is what Perlich calls the action when a customer buys a product or service after having seen an ad online. This is where Big Data comes into play – with a machine that analyzes user data and makes, on average, 50 billion decisions a day. At first, Perlich’s team looks at historical data and creates models from that information, which is the basis for the machine’s future predictions. Then, data composed of a couple of million data points, such as URL histories and app usage, feeds the computer that learns through algorithms. This helps predict whether or not a person is interested in buying a product or service. What makes digital advertising especially complicated is the need for ultrafast, automated decisions as computers must decide within 100 milliseconds which user should see what advertisement. Odds are against the statistics here because as paradoxical as it may sound, there is still not enough data available about particular user preferences. This is why, according to Perlich, advertisers do not get a good picture of what people look like in various markets. Also, data is constantly changing and there is a lot of dynamic stress on the system. But even though it is a very difficult task to get good results, Perlich believes that it works much better than the alternative of manually targeting audiences without the use of machines. She also sees the advantage in the autonomous nature of a computer’s work, which does not understand what it is doing and therefore ensures some sense of privacy for the customer. Lastly, Perlich spoke about fraud detection and the role of computer bots in the digital advertising industry.

A discussion then ensued between the panelists and the audience addressing key topics, such as user privacy and the role of computer algorithms with respect to agnostic predictive modeling. Other questions included whether automated data analysis increases inequality in society, how to define meaningful data use, and ways in which to use energy smarter. Basic research efforts at TU Dortmund aim to reduce power consumption and increase device efficiency, according to Morik. Data science works best when it is not even noticeable, Perlich said, expressing her desire to see data science being used where it matters most, such as in medicine. The panelists also highlighted a project in various African villages where satellite data is being used to distinguish stone roofs from straw roofs to prioritize the dispersal of development aid. In the future, Kersting hopes that people will become more educated about the Big Data around them in a world in which it will play an increasingly important role.

Quelle: Deutsches Wissenschafts- und Innovationshaus (DWIH) New York

Redaktion: 08.04.2016 von , Deutsches Wissenschafts- und Innovationshaus New Y

Länder / Organisationen: EU, Global, USA

Themen: Information u. Kommunikation, Ethik, Recht, Gesellschaft

[Zurück](#)

Weitere Informationen