Big Data and Digitalization: From data to understanding, insights, and knowledge

Modern information technology, incremental computing power, and online digitalization have opened up new options for utilizing automatically collected and stored data from various sources. Alongside with the continuous automated data collection there are growing needs to combine and utilize simultaneously the various data sources. As a result, there is a clear shift from the traditional computational pattern discovery of data mining to data driven modeling and inferences. Thereby Big Data is opening new potential for scientific and statistical analyses. Does this imply redefining statistical modeling and approaches at large? Does it promote descriptive statistics to a new level and demand us for new data reading abilities to reach the understanding, insights, and knowledge provided by the data?

The Graduate School of the University of Vaasa organizes a seminar in Big Data and Digitalization, and wishes welcome all researchers interested in the subject. The purpose of the seminar is to provide clarify the meaning of Big Data and present the cases as well as provide a forum for discussion. At the same time the purpose is also to assess the demands for resources and research facilities in the field at the university. More information for the doctoral students of the University of Vaasa: www.uva.fi/fi/research/researcher/graduate_school/events/bigdata/
The seminar is free of charge.

To whom: All the researchers of the University of Vaasa, Åbo Akademi, Hanken and Umeå University.
Where: University of Vaasa, Fabriikki building F119, Yliopistonranta 10
When: 9.3.2017, Thursday at 12.15–16.30
Registration: https://eforms.uwasa.fi/lomakkeet/4533/lomake.html, 6.3.2017 at the latest.

Programme

12.15–12.30 Opening of the seminar, Senior advisor **Suvi Ronkainen**

12.30–13.15 Professor **Henry Schildt**, Aalto University: *How big data and analytics are shaping strategy and management*

"Big data and sophisticated algorithms enable software to handle increasingly complex tasks. Beyond technical tasks, such as detecting fraud, optimising logistics routes, and driving cars, algorithms enable new ways to organise and control business processes. In this talk I will examine how the use of data is changing strategy processes, discuss the key impediments to data-driven management, and outline several threats posed by data-centric management practices."

13.15–14.00 Assistant professor **Petri Välisuo**, University of Vaasa: *Big measurements and data analysis, challenges and some ideas*

"Today we have better possibilities to acquire measurement data than ever before. As a consequence we sometimes have enormous amount of data even before having a well-structured hypothesis. The task of the researcher is then to make the data to tell the story about the observed reality. The challenges with big data start from storage, transfer and CPU power limitations, extending to heavy speed and dimensionality requirements of the algorithms. The work with multivariate NIR spectroscopy and with 10 TB wind turbine noise data have shown the need for choosing different approaches when the amount of data is big, such as faster algorithms, simpler models and linearisation. The resources of CSC help acquiring more CPU-power. Many challenges remain and need to be worked around."

14.00–14.20 Discussion and coffee

14.20–16.00 Big data in research – cases

Statistician **Priyantha Wijayatunga**, Umeå University: *Big data and causality – unrealistic correlation?*

"Often Big Data are only observational, but not arising from direct experiments. However, most of our research questions are causal, such as finding an effect of the social media usage on the stock market movements. In Big Data there are many instances of spurious correlations and often data-driven analyses can mislead us to false conclusions. I will talk about some of the inference problems that may arise in Big Data analytics

such as misled regressions, Simpson's paradox, false correlations, problems of adjusting for confounding and causal inference, etc."

Professor **Annukka Jokipii**, University of Vaasa: *Big data in accounting*

"In this talk I will discuss opportunities, threats and restrictions especially for accounting and auditing generated by this way of understanding information and reporting by the firms."

Assistant professor **Jukka Sihvonen**, University of Vaasa: *Quantifying Qualitative Data – Big Time*

"A major hindrance in applying big data approach to qualitative research is that the data are usually hard to collect and non-numerical. In this presentation I review the potential of using Application Programming Interfaces (APIs) in obtaining, refining and analyzing qualitative information."

16.00–16.30   Closing session and discussion,
professor **Seppo Pynnönen**