Date	**Subject**
21 June 2018	PI.lab Quarterly event Machine Learning

PI.lab Quarterly event

Prospects and limitation of responsible machine learning

Date: Friday 29 June 2018

Time: 13.00 – 16.00 (drinks afterwards)

Location: Huygensgebouw Nijmegen (room HG00.108 – ground floor)

Registration: send an email to Ms Irma Haerkens, info@pilab.nl.

The first thematically oriented quarterly PI.lab event will be dedicated to Responsible Machine Learning. Three PI.lab researchers will present their view on the topic. The presentations will be used to investigate the role the PI.lab could play or the items the PI.lab could/should address.

Programme:

13.00 – 13.10: Introduction to Quarterly Event and theme

13.10 – 13.50: Presentation and discussion **Dr. Cor Veenman** ('Responsible Data Science')

14.00 – 14.40: Presentation and discussion **Prof.dr. Arjen de Vries** ('Quantifying Bias in Search Engines and Beyond')

14.50 – 15.30: Presentation and discussion **Emre Bayamlioglu** ('A Rule based Modelling of Automated Decisions')

15.30 – 16.00: General discussion

16.00 - …: Drinks

Responsible Data Science (Cor Veenman)

"Data science is an important enabler of innovation both for commercial application as well as for improving governmental processes. With the availability of increasingly many open and closed data sources, the potential effectiveness of these data science developments grow accordingly. A great deal of the data is, however, personal and sensitive data that may be exploited in data science applications only under strict conditions. In the TNO Responsible Data Science team, we deal with pro-active developments towards data science applications that respect data protection laws and ethics in a transparent and evidence-based manner. The talk will overview data science challenges when dealing with personal data, such as detecting and suppressing protected and identifiable attributes, and detecting and repairing historical bias in the learning data."

Quantifying Bias in Search Engines and Beyond (Arjen de Vries)

"Our inherent human tendency of favoring one thing or opinion over another is reflected in every aspect of our lives, creating both latent and overt biases toward everything we see, hear, and do. Any remedy for bias must start with awareness that bias exists; for example, most mature societies raise awareness of social bias through affirmative-action programs, and, while awareness alone does not completely alleviate the problem, it helps guide us toward a solution. Bias on the Web reflects both societal and internal biases within ourselves, emerging in subtler ways. This article aims to increase awareness of the potential effects imposed on us all through bias present in Web use and content. We must thus consider and account for it in the design of Web systems that truly address people's needs."

(Intro taken from Ricardo Baeza Yates' contribution to CACM, "Bias on the Web":

https://cacm.acm.org/magazines/2018/6/228035-bias-on-the-web/fulltext).

Please, read this text. The paper is accompanied by a video, https://vimeo.com/266735243

A Rule based modelling of Automated Decisions (AD) (Emre Bayamlioglu)

"As decision-making systems are goal-oriented, their behaviour may be attributed to the inherent values and assumptions guiding their response to a given input. This allows us to infer certain normativity from the system's output as aiming to achieve some pre-set goals. Hence with normativity, we not only refer to the capacity to control and guide conduct but also to a claim, or contention, to do so which is ultimately reducible to some moral ground—say, a right to rule. Since, by themselves, facts (data) cannot provide "reasons for action", looking from the lens of normativity informs us about the motives, assumptions and the further decisional criteria underlying the systems, and thus, opens the way to a normative evaluation of the observed behaviour/action."

For the full text, see attachment.