Habilitation Thesis Reviewer´s Report

Masaryk University Faculty:	Faculty of Informatics
Procedure field:	Informatics
Applicant:	Mouzhi Ge, Ph.D.
Applicant´s home unit, institution:	Faculty of Informatics, Masaryk University
Habiltation thesis:	Data Quality Management for Recommender Systems
Reviewer:	doc. Ing. Viera Rozinajová, PhD.
Reviewer´shome unit, institution:	Faculty of Informatics and Information Technologies
	Slovak University of Technology, Slovak Republic

The habilitation thesis is written as a collection of previously published scholarly works with commentary. It interconnects two areas – recommender systems and data quality management. The topicality of given fields is high, it currently rises also in relation to ubiquitous data generation and the effort to get the most value from these data.

From among relatively rich publishing activities of the author, he selected 10 papers (4 journal papers, 4 conference papers and 2 Springer book chapters) for this thesis. Six publications fall under internationally recognized Core ranking A or B, two journals are with impact factor, one is from Current Contents database.

Comments to the **structure and the content of the thesis**: Generally, data quality is a significant aspect of all data processing tasks and if this topic is presented together with the recommender systems, I would expect to read about specific data quality problems concerned with recommendation. This is not the case in the given thesis.

Another comment concerns the selection and granularity of solved problems in the papers: The articles offer various topics – in recommendation systems area some of them are focused on the overview of the field (paper 1 and 5), diversity in recommendation list (paper 4), and some present domain specific recommender systems (paper 3 and 6). Similarly – in the domain of data quality management, subsequently the topics as data quality dimensions and assessment, data integration and an effect on decision-making are studied. In order to provide comprehensive view on both themes, a common outline across the topics would be appropriate.

In Paper 1 a comparison of different explanation types for recommender systems is presented. It was published in International Journal of Human-Computer Studies. The authors presented a set of guidelines of how to build explanations for recommender systems that support the user in decision making. They came to interesting findings concerning explanation types and in particular content-based tag cloud explanation. In the experimental part, they performed laboratory study on 105 participants from ten different countries. Then they validated the findings of the laboratory study by interview-based studies with 20 participants. However, in order to get more objective results, it would be worth to perform validation also by collecting quantitative measurements by monitoring the users during perfoming the given tasks.

In data quality management the author focuses on how to measure and improve data quality, how to avoid data quality pitfalls in data integration and how to reduce the data quality effects on business decision making. In recommender systems, there are several data quality aspects, which are significant: in particular data sparsity and data redundancy. There is no mention about these aspects in the submitted work. As the thesis is focused on two topics – recommendation and data quality, it would be beneficial to discuss such data quality problems, which are related to recommender systems.

Formal side of the work: some of the references at page 5 are incomplete – they miss information about publisher, etc. Some tipos are in the document and several sentences do not start with capital letter.

Concerning the **given criteria of the Faculty of Informatics of Masaryk University** – the candidate has published 4 articles in scholarly journals listed in WoS, 1 article in scholarly journal listed in Scopus, he has altogether 10 journal papers, 5 chapters in scholarly book publications and 40 publications in proceedings of conferences listed in Scopus. He also co-edited 4 scholarly published book publications or thematic proceeding.

He is very active in academic and scholarly community - he has organized more than 10 workshops at international conferences.

He has 307 citations from Scopus (self-citations are excluded).

Pedagogical activities include teaching at Technical University of Dortmund in Germany, Free University of Bozen – Bolzano in Italy and Masaryk University.

In my opinion, the reviewed proposal complies with the requirements posed on theses aimed for habilitation procedure.

Reviewer´s questions for the habilitation thesis defence:

- As you declared in the „Scholarly and Academic Activities", you plan to carry out on your research on explanation in recommender systems and user-centric evaluation for recommender systems. Could you discuss in more detail the evaluation of the explanation systems, described in Paper 1? Wouldn´t it be useful to collect more quantitative measurements by monitoring the user when perfoming the tasks?
- The data quality is important for all tasks of data analysis area. Please specify and discuss those aspects of this field, which are particularly important for recommender systems.

Conclusion:

The habilitation thesis entitled „**Data Quality Management for Recommender Systems**" by Mouzhi Ge *fulfills* requirements expected of a habilitation thesis in the field of Informatics.

In Bratislava, on 18-th May 2018   ....

                                                                                    signature