



Habilitation Thesis Reviewer's Report

Masaryk University	
Faculty	Faculty of Informatics
Procedure field	Informatics
Applicant	Mouzhi Ge
Applicant's home unit, institution	Faculty of Informatics, Masaryk University
Habilitation thesis	<i>Data Quality Management for Recommender Systems</i>
Reviewer	Doc. Ing. Přemysl Brada, MSc. Ph.D.
Reviewer's home unit, institution	Faculty of Applied Sciences, University of West Bohemia, Pilsen, CZ

The habilitation thesis of Mouzhi Ge is placed on the intersection of data quality management, decision making and recommender systems. This is an area which has seen rapid development in both research and industry, with clear impact on both businesses and people's daily life.

As presented in the thesis, the area of information and data quality is considered as a condition necessary for well-functioning recommender systems. The candidate has, among other results, proposed a framework for assessing information quality using a well-defined, measurable set of criteria suitable for practical usage. This research further led to the investigation into the relationship between information quality and decision making. The conclusion that information accuracy and completeness affect decision quality significantly is an important result with practical implications for the information management field.

Preceding in the thesis, but perhaps logically consequent, is the area of recommender systems. Here the candidate has created a contribution towards assessing the state of research in this area, with a highly cited literature review paper published at a CORE B conference plus other survey papers. The candidate's research into the diversity of recommendations and enhanced explanatory approaches, summarized in sections 2.2 and 2.3 of the thesis, can be partially traced to the findings of that review paper.

These results provide a good evidence that the candidate has established a distinct research agenda which he has been steadily pursuing and which has lead to his continuous contributions to the field. In the thesis itself the relationship between the two areas of data quality and recommender systems is explained, most prominently in the Abstract and at the end of the Introduction, forming a (tacit) motivation for that research agenda.

It can only be highly commended that the candidate has been using various forms of quantitative research approaches in many of the research works, from literature reviews to well-designed experimental evaluations of the proposed methods. On the other hand, it is unclear from the thesis whether there have been any attempts on an industrial uptake of the

candidate's research results. I would have expected this aspect to be at least mentioned as many of the presented topics are highly relevant to the industry.

Regarding the thesis as a commented collection of papers, the papers have clearly been selected carefully for the collection and are mostly of high quality. I believe nevertheless that adding some other publications would have provided a more complete picture of the contributions of the candidate in the individual topics. A wider selection of papers would also potentially have provided a better support for the interconnection between the two major areas of the thesis, as mentioned above. (One formal issue related to the list of papers: the Chapter "Data Quality Problems in TPC-DI based Data Integration Processes" (2017) does not appear to be available yet from the publisher and thus could not be verified.)

Beside the habilitation thesis itself, the candidate's research record is of a very good standing, with 5 indexed journal publication and 40 papers in the proceedings of indexed conferences. Some of the candidate's works have been frequently cited, resulting in a total of 416 Scopus citations to date. Also, the candidate has an excellent track of international cooperation.

Reviewer's questions for the habilitation thesis defence (number of questions up to the reviewer)

1. Majority of the data quality results summarised in the thesis provide new insights into the assessment of data quality. How could these insights be transformed into improvements in data management and governance methods, such that would lead to the availability of raw data – and thus the derived information – of higher quality?

Conclusion

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

In Pilsen on 16th May 2018

