Limits and Potentials of Grasping User Interaction in Controlled and Open World Environments

Contrasting PIAAC Large-Scale-Assessment Usage Data with Web Analytics

Peter Böhm, Christoph Schindler,
Heiko Rölke, Marc Rittberger

German Institute for International Educational Research, Germany
{boehm, schindler, roelke, rittberger}@dipf.de

Keywords: usage data; web analytics; log data; computer-based assessment; web portal; qualitative analysis

1 Motivation and background

Educational practices are becoming more and more digitalized and learning is realized in the open web as well as in controlled digital learning environments. This provides new opportunities not only for new interactive teaching methods, but also for tracking the interaction with learning materials for improving teaching and learning as well as for assessments (cf. Siemens & Long, 2011). On one hand, the field of Computer-based assessment (CBA) uses these new research capacities for improving competence assessments in controlled and validated test environments. On the other hand, in the open web, tasks are pursued in interactive situations and protocolled in logfiles. In this study, these two different situations of open and controlled environments are contrasted by analyzing the different fabrications and expressiveness' of

data. The main research questions are: To what degree can open and closed world log data of learning tasks be compared? Where do they differ and what conclusions can be drawn regarding the grasping of user interaction and possible mutual optimizations of the log data gathering? By addressing these questions, hitherto disparate areas of research are linked.

2   Research design

PIAAC (Programme for the International Assessment of Adult Competencies) is used as a case study for a controlled environment of a large-scale assessment with CBA tools, while the German Education Server (GES)[1] serves as a case study for the open data environment. Here, like in the field of CBA, usage data is gathered but its expressiveness is limited because of the open, uncontrolled usage and usage data (Keil et al., 2015).

This study is methodologically based on an in-depth, qualitative comparison of the quantitative usage data of one PIAAC test item and a comparable usage scenario of the GES. This comparison includes the production conditions of the test items, the usage data, the fabrication and syntax of the data as well as the expressiveness' of metrics that can be calculated based on these data.

3   Expected outcome

By comparing these two different situations of data fabrications and the expressiveness of the data, we aim to analyze the limits and potentials of usage data from closed, highly controlled environments and open-world, uncontrolled websites. With open-world usage data, difficulties are expected in reliably recognizing visitors and the 'real' browsing behavior, while large quantities of log data are available. By contrast, CBA usage data is more controlled; in the case of PIAAC even extensive background information is

---

1 See http://www.eduserver.de/.

available on the participants. The number of participants is usually smaller than the number of users of large websites, whereby the test is realized under 'artificial´ test conditions. These first insights address an exchange on an epistemological level between both worlds (CBA and web analytics), advancing the examination of expressiveness of user interaction in different learning settings.

References

Keil, S., P. Böhm and M. Rittberger (2015): Qualitative web analytics: New Insights into Navigation Analysis and User Behavior – A Case Study of the German Education Server. In: F. Pehar et al. (Eds.): *Re:inventing Information Science in the Networked Society.* Proceedings of the 14th International Symposium on Information Science (ISI 2015), Zadar, Croatia, 19th–21st May 2015 (pp. 252–263). Glückstadt: Hülsbusch.

Siemens, G. and P. Long (2011): Penetrating the Fog: Analytics in Learning and Education. In: *EDUCAUSE review*, 46 (5), 30.