

Summary of the 3rd meeting of the

***dblp* Advisory Board**

held on June 1st, 2012,
in Mannheim, Germany



SCHLOSS DAGSTUHL
Leibniz-Zentrum für Informatik

In attendance:

- Dietmar Saupe (Univ. of Konstanz)
- Hannah Bast (Univ. of Freiburg)
- Hans-Peter Lenhof (Saarland Univ.)
- Marcel R. Ackermann (Schloss Dagstuhl LZI, *dblp*)
- Michael Ley (Univ. of Trier, *dblp*)
- Michael Wagner (Schloss Dagstuhl LZI, *dblp*)
- Mila Majster-Cederbaum (Univ. of Mannheim)
- Reinhard Wilhelm (Schloss Dagstuhl LZI, Saarland Univ.)
- Rüdiger Dillmann (Karlsruhe Inst. of Tech.)
- Rüdiger Reischuk (Univ. of Lübeck)

Absent (excused):

- Andreas Butz (LMU Munich)
- Oliver Günther (Univ. of Potsdam)
- Jürgen Teich (Univ. of Erlangen-Nuremberg)
- Otto Spaniol (RWTH Aachen)

Agenda

1. Opening remarks and *dblp* progress report
2. Bibliometric survey
3. *dblp* venue application
4. Fake conferences
5. Perspectives and further development
6. Appointment of next meeting

Meeting begins: 1:00 p.m.

Item 1: Opening remarks and *dblp* progress report

Hannah Bast welcomes the board members and opens the meeting.

Marcel R. Ackermann gives a brief overview on the recent progress of *dblp* (c.f. the accompanying slides). The key developments are:

- *dblp* is about to hit the 2 million publications milestone in early June 2012
 - over 6.000 (mainly German) dissertations have been added
 - many back-issues of CORE/CAPES A* venues have been added
 - venues suggested by the *dblp* Advisory Board has been added
- New *dblp* logo has been designed
- Redesign of *dblp* website is in progress
- Initiated colloquium series “computer science and bibliometrics” at Saarland University
- DFG project “Smart Harvesting” has been granted (Univ. o. Trier together with GESIS)

Item 2: Bibliometric survey

Michael Wagner presents the results from a number of small surveys that took place at Dagstuhl seminars. The survey asks for the top 5 most relevant conference and the top 5 most relevant journals in the area of expertise of seminar participants. The results from the surveys have been used to identify relevant venues that are missing in *dblp*.

Michael Wagner also presents the prototype for the planned large-scale survey. The web based survey system is demonstrated. As in the Dagstuhl survey, the prototype asks for up to five conferences and journals each. It also asks the expert to classify each venue by a field of research classification scheme.

It is discussed whether the current design of the survey allows for *dblp* to learn of new conferences that *dblp* does not already know about. It is conjectured that by asking only for the top 5 venues, answers will center around the same small set of known top venues, and many interesting venues in the “heavy tail” of possible answers might be neglected. A number of possible improvements to the survey are discussed:

- The expert could be allowed to provide information on as many venues as she sees fit (instead of only 5). In doing so, the expert would be encouraged to provide a more wholesome picture of her field of expertise. It is noted that with an arbitrary long list of venues, and additional questions for each given venue (e.g., asking for the venue’s field of research), this increases the time necessary for the expert to complete the survey and might reduce the expert’s willingness to fill out the survey completely.
- After giving her list of venues, the expert could also be queried for a small number of (randomly chosen) venues from her field of expertise that have not yet received many ratings. Thereby, information for neglected venues could be obtained. A technical problem of this approach is that the classification by fields of expertise does not exist for most venues beforehand.
- After a certain number of surveys have been collected, the survey scheme could adapt to the already given answers to direct the next expert to give new answers. E.g., the expert could be asked for her most relevant venues besides a given list of venues for which plenty of answers exist. One downside of this feedback loop is that the assumption of statistical independence of the given answers is no longer valid. This will render some basic analysis techniques such as averaging or quantitative rankings useless.

It is noted that some of the suggestions may be infeasible with the current choice of survey software.

The importance of a sound field of research classification scheme is discussed. Michael Wagner has

prepared an overview of existing classification schemes, but concludes that they are unsuitable for different reasons (too few/many categories, imbalanced category sizes, important fields missing, etc.).

Michael Wagner also presents an own proposal of a field of research classification scheme. The scheme is briefly discussed.

Actions:

- Michael Wagner (in coordination with Jürgen Teich and Rüdiger Reischuk) is asked to evaluate the possible improvements to the survey system and provide some prototypical implementations. The discussion should be continued via email.
- Michael Wagner will contact the GI to investigate whether GI members are available as experts.
- Each member of the board is encouraged to evaluate the proposed field of research classification scheme and to point out possible flaws. The discussion should be continued via email.

Item 3: *dblp* venue application

Marcel R. Ackermann presents the proposed venue application policy. The policy is discussed and a number of small corrections/additions are made. The updated policy is passed by the board.

Actions:

- *dblp* will select some recent applications and use them as test cases for the application process.
- Marcel R. Ackermann will set up the JIRA web platform as soon as possible such that applications can be accessed online.

Item 4: Fake conferences

It is discussed how *dblp* should handle fake conferences or fake journals. A number of dubious venues that are indexed in *dblp* are inspected. It is discussed that it is hard to distinguish malicious fake venues from good-natured but low-quality venues. There is no reliable index of malicious fake venues known.

Consensus is reached that the new venue application policy will serve as a safety measure against fake venues. *dblp* should stop indexing a venue whenever malicious practices come to our attention. When in doubt, the venue should be asked to undergo the venue review process.

There was agreement that existing listings even from (by now) well-known fake conferences should not be removed, mainly in order to avoid (legal) confrontations with the people involved.

Marcel R. Ackermann remarks that *dblp* invited Prof. Debora Weber-Wulff to talk on fake conferences at the “computer science and bibliometrics” colloquium series on July 16, 2012 (at Saarland University).

Item 5: Perspectives and further development

Marcel R. Ackermann gives a demonstration of some new features that are currently in development at *dblp*. These features include:

- the *dblp* error report portal,
- institution data for authors,
- subject tags for venues,
- the new *dblp* web page.

It is discussed whether the default order of the new *dblp* author web page should be chronologically or

by publication type. Marcel R. Ackermann reports that the choice of the publication type default order is intentional to highlight book and journal publications, and to reduce the prominence of informal publications such as documents on preprint servers. The board suggests to use the chronological order as default order and to avoid such a “political” statement.

Item 6: Appointment of next meeting

The next meeting will take place preferably on November 16th, 2012, in Mannheim, Germany. Mila Majster-Cederbaum volunteered to try to arrange the chancelor’s meeting room again. This is highly appreciated by all participants.

Meeting adjourns: 4:30 p.m.

Meeting notes: Marcel R. Ackermann, Trier, Germany, June 25th, 2012