

# LZI+DBLP

Zwischenbericht Mai 2012

Marcel R. Ackermann\*    Marc Herbstritt\*    Oliver Hoffmann\*  
Michael Ley†    Christian Lindig\*    Michael Wagner\*

## Zusammenfassung

Die Informatik benötigt eine belastbare Datenbasis zum Nachweis und zur Evaluierung wissenschaftlicher Literatur. In dem von der Leibniz-Gemeinschaft geförderten Projekt „LZI+DBLP“ entwickeln Schloss Dagstuhl – Leibniz-Zentrum für Informatik und die Bibliographiedatenbank *dblp* gemeinsam die dafür notwendigen technischen, inhaltlichen und organisatorischen Strukturen. Das Projekt wird zudem durch die Klaus Tschira Stiftung gefördert.

Dieser Zwischenbericht dokumentiert den Stand des Projektes ein Jahr nach offiziellem Projektbeginn. Er zeigt, in welchem Umfang im Laufe des Projektes die Datenbasis von *dblp* erweitert und vertieft werden konnte. Zudem berichtet er vom Aufbau neuer organisatorischer Strukturen und Maßnahmen zur Sicherung der hohen Datenqualität von *dblp*.

## 1 Einleitung

Die Evaluierung wissenschaftlicher Literatur erfordert eine belastbare Datenbasis für wissenschaftliche Publikationen, der eine nachhaltige Infrastruktur zugrunde liegt. Die einzigartige Publikationskultur in der Informatik, die einen starken Schwerpunkt auf Konferenzpublikationen legt, stellt dabei besondere Anforderungen, denen allgemeinwissenschaftliche Literaturdatenbanken in der Regel nicht gerecht werden. Schloss Dagstuhl – Leibniz-Zentrum für Informatik (LZI) und die Universität Trier arbeiten seit Anfang 2010 zusammen mit dem gemeinsamen Ziel, die Bibliographiedatenbank *dblp* gemäß ihrer nationalen und internationalen Bedeutung als zentralen Nachweis für wissenschaftliche Literatur in der Informatik und der Informatik nahestehenden Gebieten inhaltlich, organisatorisch und technisch zu stärken und auszubauen.

---

\*Schloss Dagstuhl – Leibniz-Zentrum für Informatik, <http://www.dagstuhl.de>

†dblp computer science bibliography, Universität Trier, <http://www.dblp.org/db/>

Zu diesem Zweck wurde im Juni 2011 das über den Pakt für Forschung und Innovation der Leibniz-Gemeinschaft geförderte Projekt „LZI+DBLP: Konsolidierung der bibliometrischen Datenbasis in der Informatik“ offiziell gestartet. Dieser Zwischenbericht dient zur Dokumentation der erzielten Ergebnisse des Projektzeitraumes bis Mai 2012. Die Ergebnisse der Projektvorbereitungsphase von November 2010 bis April 2011 können dem Zwischenbericht aus dem vergangenen Jahr entnommen werden [2].

## 2 Hintergrund und organisatorische Struktur

Die Datenbank *dblp* besteht seit 1993 und wird seitdem von ihrem Gründer Dr. Michael Ley gepflegt und erweitert. Zunächst als Ein-Personen-Projekt gestartet und später durch gelegentliche Spendenmittel, wissenschaftliche Mitarbeiter und studentische Hilfskräfte unterstützt, entwickelte sich *dblp* über die Jahre zu der internationalen Referenzdatenbank bibliographischer Metadaten in der Informatik. Die Personalstärke von *dblp* blieb dabei jedoch stets deutlich unterhalb der gemäß Aufgabe und Bedeutung angemessenen Notwendigkeit.

Um den Nutzen von *dblp* für die nationale und internationale Informatik-Forschung langfristig zu konsolidieren und zu stärken wurde Anfang 2010 eine Zusammenarbeit zwischen *dblp* und Schloss Dagstuhl – Leibniz-Zentrum für Informatik begonnen. Der wissenschaftliche Beirat, das industrielle Kuratorium und das wissenschaftliche Direktorium von Schloss Dagstuhl sprachen sich dabei nachdrücklich für eine Kooperation zwischen Schloss Dagstuhl und *dblp* aus. Ebenso positiv war die Resonanz durch den Fakultätentag Informatik<sup>1</sup> und den Beirat der Universitätsprofessoren der Gesellschaft für Informatik (GIBU)<sup>2</sup>, denen die Zusammenarbeit 2010 vorgestellt wurde.

Schloss Dagstuhl hat deshalb im März 2010 beim Senatsausschuss Wettbewerb der Leibniz-Gemeinschaft<sup>3</sup> einen Antrag auf Förderung gestellt, der im Dezember 2010 bewilligt wurde und im Wesentlichen die Finanzierung von zwei wissenschaftlichen Stellen für die Dauer von zwei Jahren vorsieht.

Dank der Finanzierung war es möglich, ein Team unter der Koordination von Dr. Michael Ley an der Universität Trier zu formieren. Mit zunächst einer durch Dipl.-Inform. Oliver Hoffmann besetzten Stelle begann das Projekt offiziell im Juni 2011 und wurde ab August 2011 durch Dr. Marcel R. Ackermann verstärkt. Die Datenerfassung und Qualitätskontrolle wird zudem seit Juli 2011 halbtags von Frau Stefanie von Keutz unterstützt.

---

<sup>1</sup><http://www.ft-informatik.de/>

<sup>2</sup><http://www.gibu.gi-ev.de/>

<sup>3</sup><http://www.wgl.de/?nid=wet>

Bereits vorab wurde das Vorhaben durch eine Spende in Höhe von 25.000 € von der Klaus Tschira Stiftung gefördert. Die Spende ging im November 2010 ein und diente der Finanzierung der Stelle von Herrn Oliver Hoffmann ab diesem Zeitpunkt. Dadurch konnten Vorarbeiten für das Projekt bereits vor dem offiziellen Projektbeginn initiiert werden.

Dank einer weiteren Spende der Klaus Tschira Stiftung in Höhe von insgesamt 120.000 € konnte im Januar 2012 mit Dipl.-Inform. Michael Wagner eine weitere Vollzeitkraft zunächst für zwei Jahre für das Projekt gewonnen werden.

Seit März 2012 wird das *dblp*-Team bei der Administration der *dblp*-Server von Herrn Christopher Perrin (studentische Hilfskraft) unterstützt.

### 3 Projektfortschritt Mai 2011 – Mai 2012

#### 3.1 Wissenschaftliche Aufsicht

Eine wesentliche Neuerung des vergangenen Projektjahres ist die Formierung eines Beirates unter dem Dach von Schloss Dagstuhl, welches die strategischen Entscheidungen hinsichtlich der Entwicklung von *dblp* verantwortet. Zu den Aufgaben des *dblp Advisory Boards* gehören insbesondere:

- Festlegung von Richtlinien, Qualitäts- und Prioritätsvorgaben zur inhaltlichen Ausrichtung von *dblp*,
- wissenschaftliche Aufsicht über die inhaltliche Qualität des Datenbestandes,
- ideelle Förderung von *dblp* durch das Einbringen von Fachkompetenz und
- Repräsentation von *dblp* gegenüber der wissenschaftlichen Gemeinschaft.

Dabei ist es gelungen, für das Board zehn namhafte Persönlichkeiten aus verschiedenen Disziplinen der Informatik zu gewinnen. Dem *dblp Advisory Board* gehören an:

- Prof. Dr. Andreas Butz (LMU München, *Media Informatics and Human-Machine-Interaction*),
- Prof. Dr. Dietmar Saupe (Universität Konstanz, *Multimedia Signal Processing*),
- Prof. Dr. Hannah Bast (Universität Freiburg, *Algorithms and Data Structures*),
- Prof. Dr. Hans-Peter Lenhof (Universität des Saarlandes, *Bioinformatics*),

- Prof. Dr.-Ing. Jürgen Teich (Universität Erlangen-Nürnberg, *Hardware-Software-Codesign*),
- Prof. Dr. Mila Majster-Cederbaum (Universität Mannheim, *Complex Systems*),
- Prof. Oliver Günther, Ph.D. (Universität Potsdam, *Information Systems*),
- Prof. Dr. Dr. h.c. Otto Spaniol (RWTH Aachen, *Communication and Distributed Systems*),
- Prof. Dr.-Ing. Rüdiger Dillmann (Karlsruher Institut für Technologie, *Humanoids and Intelligence Systems*),
- Prof. Dr. Rüdiger Reischuk (Universität zu Lübeck, *Theoretical Computer Science*).

Das *dblp* Advisory Board hat sich im November 2011 in Saarbrücken konstituiert. Sprecher des Boards ist Hannah Bast. Zu den ersten Zielen des Boards zählen die Verabschiedung von Mindestanforderungen und Rahmenrichtlinien für die Aufnahme neuer Publikationsreihen in *dblp*, sowie die Definition von Anforderungen für den Aufbau einer bibliometrischen Infrastruktur.

### 3.2 Technische Weiterentwicklung

Die Literaturdatenbank *dblp* verzeichnet wissenschaftliche Literatur in der Informatik auf der Ebene einzelner Beiträge. Sie konzentriert sich bei der Wahl der Datenquellen auf Verlage und Bibliotheken und verfolgt das Ziel, ganze Reihen und Serien und damit Themengebiete wissenschaftlicher Literatur möglichst vollständig zu erfassen.

Über Jahre hinweg wurden diese Daten stets mit hohem manuellen Anteil erfasst. Dies führte zu einer anerkannt hohen Qualität des Datenbestandes, stellt aber einen hohen Arbeitsaufwand dar und kann der sich stetig ausweitenden Publikationslandschaft nicht adäquat Rechnung tragen. Bereits in der Projektvorbereitungsphase wurde daher eine spezialisierte Software, so genannte *Wrapper*, entwickelt, die bibliographische Rohdaten von Web-Seiten wissenschaftlicher Verlage und Bibliotheken sammelt [1]. Dies entspricht dem Vorgehen einer Suchmaschine wie Google, die durch *Crawler* Web-Seiten automatisch besucht und zur Indexierung vorbereitet.

Diese Software wurde im Rahmen des Projektes stetig erweitert und ausgebaut. Derzeit (Stand Ende Mai 2012) kommen Wrapper bei 89 verschiedenen digitalen Bibliotheken und Verlagssystemen zum Einsatz und stellen heute einen integralen Teil des Arbeitsablaufes bei der Neuaufnahme von Daten dar. Die abschließende Qualitätskontrolle und Fehlerbeseitigung hingegen wird zwar von Hilfssoftware unterstützt; die Endabnahme erfolgt jedoch auch weiterhin bewusst von Hand.

In einem gemeinsamen Projektvorhaben von *dblp* und GESIS – Leibniz-Institut für Sozialwissenschaften soll zudem die Schaffung von Werkzeugen und Verfahren für das wissenschaftlichen Informationsmanagements erforscht werden. Ziel des Vorhabens ist es unter Anderem, die bisher rein regelbasierte Wrapper-Software durch generische, adaptive oder lernende Verfahren zu erweitern, um diese noch breiter einsetzen zu können. Von solchen „smart harvesting“ Verfahren würde *dblp* in besonderem Maße profitieren. Der Projektantrag wurde im April 2012 vom Senat der DFG bewilligt; als Projektbeginn wird das vierte Quartal 2012 anvisiert.

Des Weiteren wurde unter der Adresse [dblp.dagstuhl.de](http://dblp.dagstuhl.de) ein neuer, leistungsfähigerer *dblp*-Server in Schloss Dagstuhl in Betrieb genommen. Bei der Administration arbeitet das *dblp*-Team mit der IT-Abteilung von Schloss Dagstuhl unter der Leitung von Dipl.-Inform. Thomas Schillo zusammen. Der neue Server fungiert derzeit noch als Mirror der bestehenden *dblp*-Webstruktur, soll aber in naher Zukunft unter der Domain [dblp.org](http://dblp.org) zum zentralen *dblp*-Server ausgebaut werden.

### 3.3 Produktivität der Datenakquise

Bei Beginn der Zusammenarbeit von Schloss Dagstuhl und *dblp*, Anfang November 2010, indexierte *dblp* etwa 1,49 Millionen Publikationen. Ende Mai 2012 ist *dblp* um über ein Drittel gewachsen und indexiert nun bereits 1,98 Millionen Publikationen. Details der Produktivitätsentwicklung von *dblp* werden in Tabelle 1 zusammengefasst.

Im Vergleich zum Vorjahr wurde die Produktivität bei der Erfassung neuer Datensätze abermals gesteigert. Im Vergleichszeitraum vor Projektbeginn, in den sechs Monaten von November 2009 bis April 2010, wurden 96.658 neue Publikationen in *dblp* aufgenommen. In der Vorbereitungsphase von November 2010 bis April 2011 konnte dies bereits auf 156.390 neue Publikationseinträge gesteigert werden. Im Zeitraum von November 2011 bis April 2012 konnte die Anzahl der Neuaufnahmen erneut auf 178.643 Neueinträge gesteigert werden. Für das kommende Projektjahr gehen wir davon aus, diese Produktivität halten zu können. Die im Projektantrag angepeilte Erfassungsquote von 200.000 Datensätzen pro Jahr dürfte damit auf jeden Fall mehr als erfüllt werden.

### 3.4 Aktualität neuer Einträge

Die Literaturdatenbank *dblp* soll neue Literatur möglichst schnell und vollständig erfassen, um die große Nachfrage insbesondere bei der Recherche

Anzahl neuer Publikationen in <i>dblp</i>										
Nov. 2009 bis Apr. 2010		Mai 2010 bis Okt. 2010		Nov. 2010 bis Apr. 2011		Mai 2011 bis Okt. 2011		Nov. 2011 bis Apr. 2012		
Anzahl	Jahr	Anzahl	Jahr	Anzahl	Jahr	Anzahl	Jahr	Anzahl	Jahr	Alter
0	neuer	544	neuer	1	neuer	704	neuer	0	neuer	zukünftig
16753	2010	59079	2010	26154	2011	87869	2011	38078	2012	aktuell
55830	2009	25333	2009	68647	2010	21557	2010	68107	2011	-1 Jahr
9564	2008	6211	2008	18441	2009	8733	2009	15181	2010	-2 Jahre
3269	2007	2979	2007	6911	2008	5064	2008	9512	2009	-3 Jahre
11483	älter	8422	älter	36114	älter	19064	älter	47765	älter	älter
96899	total	102568	total	156268	total	142991	total	178643	total	

**Tabelle 1:** Produktivitätsentwicklung *dblp*. Die Tabelle gibt eine Übersicht über die Anzahl neu aufgenommenen Publikationseinträge in den Sechmonatsintervallen vor und seit November 2010. Zu jedem Intervall sind die Neuaufnahmen nach ihrem jeweiligen offiziellen Publikationsjahr aufgeschlüsselt. Es ist zu beachten, dass auf Grund der *online first* Strategie einiger Verleger in *dblp* auch Publikationen gelistet sind, deren Publikationsdatum erst in einem kommenden Jahr liegt. Der nur scheinbare Produktivitätsrückgang zwischen Mai 2011 und Oktober 2011 ist durch saisonale Schwankungen im Publikationskalender und die Urlaubszeiten im Sommer begründet.

von aktuellen Publikationen befriedigen zu können. Ein Maß für die Aktualität einer Veröffentlichung ist das Alter bei ihrer Erfassung, d.h., der Zeitraum zwischen dem Erscheinen der Veröffentlichung und ihrem Eintrag in *dblp*.

Um die Entwicklung der Aktualität zu messen, haben wir das Alter von neu erfassten Veröffentlichungen verglichen. Tabelle 1 fasst diese Ergebnisse zusammen. Dabei zeigt sich, dass zwischen November 2011 und April 2012 die Aufnahme aktueller Publikationen gegenüber dem Vergleichszeitraum des Vorjahres noch einmal deutlich gesteigert werden konnte, ohne zu Lasten der Vervollständigung älterer Publikationsreihen zu gehen. So materialisierte sich in der Steigerung von 26.154 auf 38.078 Publikationen bei der Neuaufnahme jahresaktueller Publikationen etwa die Hälfte der Produktivitätssteigerung des ersten Projektjahres. Die andere Hälfte ist der Aufnahme von Publikation mit einem Publikationsdatum von vor 3 Jahren oder älter zu Gute gekommen (vgl. Abschnitt 3.6).

### 3.5 Neuaufnahme von Publikationsreihen

Trotz der hohen Reputation von *dblp* ist sowohl die thematische Breite in den Randgebieten der Informatik als auch die Abdeckung einzelner Themengebiete der Kern-Informatik noch unzureichend. In der Vergangenheit führten vor allem fehlende Ressourcen dazu, dass viele Publikationsreihen nicht

berücksichtigt werden können. Als Auswahlkriterien zur Aufnahme neuer Reihen wurden lange Zeit vor allem die eigene Erfahrung sowie persönliche Kontakte herangezogen. Diesen Prozess galt es in der ersten Projektphase auf ein solideres Fundament zu stellen.

Da *dblp* bisher noch keine eigene Infrastruktur zur Bewertung von Publikationsreihen etabliert hat, wurde im ersten Projektjahr verstärkt externe Sachkompetenz herangezogen. Zum einen konnten mit dem neu konstituierten *dblp* Advisory Board erstmals eine Runde von Experten aus verschiedenen Disziplinen der Informatik direkt ihre Expertise in den Auswahlprozess einbringen. Mitglieder des Boards beraten das *dblp*-Team bei Entscheidungsfragen und weisen zudem aktiv auf Fehlstände im *dblp*-Katalog hin. Darüber hinaus konnten in vielen Gesprächen mit Teilnehmern von Dagstuhl-Seminaren Einschätzungen und Empfehlungen von anerkannten Fachexperten gewonnen werden. In diesem Austausch offenbarte sich ein weiterer, großer Vorteil der Kooperation von *dblp* und Schloss Dagstuhl.

Neben der Integration von Expertenmeinungen in den Auswahlprozess haben sich Entscheidungen zudem an anerkannten Informatik-Rankings orientiert. Hier sind vor allem das CORE-Ranking<sup>4</sup> der *Computing Research and Education Association of Australasia* sowie das CAPES-Ranking<sup>5</sup> des brasilianischen Forschungsministeriums hervorzuheben. Beide Rankings bieten — wenn gleich auch regional geprägt — eine wertvolle Hilfestellung bei der Suche nach wichtigen Publikationsreihen.

Auf diesem Wege konnten im ersten Projektjahr 389 neue Reihen identifiziert und aufgenommen werden. Die Anzahl neu aufgenommener Reihen kann Tabelle 2 entnommen werden.

### 3.6 Vollständigkeit bestehender Reihen

Neben der Aufnahme neuer Publikationsreihen liegt ein weiterer Schwerpunkt auf der Vervollständigung bereits indexierter Zeitschriften und Konferenzserien. Hier wurden im vergangenen Projektjahr verstärkt Anstrengungen unternommen, insbesondere die wissenschaftlich bedeutendsten Reihen lückenlos in *dblp* zu indexieren.

Als Auswahlkriterien wurden dabei ebenfalls die etablierten CORE- und CAPES-Rankings herangezogen. Fokus der Bemühungen war dabei insbesondere, die Reihen der Spitzenkategorien vollständig, sowie hochklassige Reihen zumindest bis in das Jahr 2000 vollständig zu ergänzen. Hier konnten

---

<sup>4</sup><http://core.edu.au/>

<sup>5</sup><http://qualis.capes.gov.br/webqualis/>

Anzahl Publikationen/Reihen in <i>dblp</i>						
	Journale			Konferenzen/Workshops		
	Reihen	Ausgaben	Artikel	Reihen	Bände	Papiere
Apr. 2010	820	13436	534727	2794	15388	818707
Okt. 2010	852	14031	577743	2877	16184	877320
Apr. 2011	1044	15874	664129	2950	16939	937194
Okt. 2011	1091	16591	722161	3094	18138	1019094
Apr. 2012	1179	18139	806739	3204	19311	1100778

**Tabelle 2:** Journale und Konferenz-/Workshopreihen in *dblp*. Die Tabelle beschreibt die Trendentwicklung bei der Anzahl von Publikationsreihen, Ausgaben und Artikel in *dblp*. Es ist zu beachten, dass die Aufteilung in Reihen und Bände insbesondere bei Konferenzen und Workshops nicht immer eindeutig ist. Die vorliegende Zählung betrachtet daher die Aufteilung in Übersichtsseiten (Reihen) und Inhaltsverzeichnisse (Bände), wie sie auf der *dblp*-Webseite abgebildet ist, und spiegelt dabei die allgemeine Tendenz wieder.

inzwischen große Lücken in der Abdeckung geschlossen werden. Die Vervollständigung aller prestigeträchtigen Reihen ist dabei ein Prozess, der noch immer andauert.

### 3.7 Neuaufnahme von Monographien

Ein weiteres Ziel des Projektes ist die Verbreiterung der Datenbasis im Bereich von Monographien und Sammelbänden. Hierbei wurde im ersten Projektjahr die Zusammenarbeit mit der Informatik-Fachbibliothek von Schloss Dagstuhl gepflegt und intensiviert. Nachdem im Zeitraum vor Projektbeginn zunächst die Integration des Ist-Bestandes der Bibliothek im Vordergrund stand wurde seitdem ein regelmäßiger Datenaustausch zwischen Bibliothek und *dblp* installiert, bei dem die monatlichen Bibliotheks-Neuzugänge weitgehend in *dblp* integriert werden. Insgesamt wurden so zwischen Mai 2011 und April 2012 insgesamt 1.779 neue Monographien in den Datenbestand übernommen. Die konkreten Zahlen können Tabelle 3 entnommen werden.

Ein weiterer Schwerpunkt bildet die Neuaufnahme von Dissertationen. Hier konnte *dblp* im ersten Projektjahr um insgesamt 6.090 Dissertationen von zumeist deutschen Hochschulen erweitert werden. Es ist ein Ziel von *dblp*, die Aufnahme von Dissertationen zu intensivieren und zudem auf den internationalen Bereich auszudehnen.



	Anzahl Monographien in <i>dblp</i>			
	Bücher/Kollektionen		Dissertationen	
	Anzahl	Zuwachs	Anzahl	Zuwachs
Apr. 2010	1600		91	
Okt. 2010	1675	+ 75	91	+ 0
Apr. 2011	7484	+ 5809	808	+ 717
Okt. 2011	8851	+ 1367	829	+ 21
Apr. 2012	9263	+ 412	6898	+ 6069

**Tabelle 3:** Monographien in *dblp*. Die Tabelle fasst die Anzahl in *dblp* indexierter Monographien im Sechsmonatsabstand zusammen. Der Sprung von Oktober 2010 zu April 2011 geht auf die Integration des Ist-Bestandes der Informatik-Fachbibliothek von Schloss Dagstuhl zurück. Der Anstieg bei den Dissertationen zum April 2012 wurde durch das gezielte Sammeln von Metadaten über die Webseiten deutscher Hochschulen und der Deutschen Nationalbibliothek erreicht.

### 3.8 Kontakte mit Verlagen als Datenlieferanten

Der Kontakt zu Verlagen und Fachgesellschaften ist notwendig, um über Neuerscheinungen informiert zu bleiben, sowie um möglichst zeitnah an bibliographische Metadaten zu gelangen. Im Idealfall erweisen sich Verlage dabei als kontinuierliche Datenlieferanten.

Bereits im Mai 2011 konnte mit *Springer Science+Business Media*<sup>6</sup> einer der für die Informatik wichtigsten Verleger als Partner und Datenlieferant gewonnen werden. Seitdem beliefert Springer *dblp* regelmäßig mit den Metadaten zu allen Neuerscheinungen. Die Daten werden tagesaktuell in den *dblp* Datenbestand eingepflegt.

Seit Oktober 2011 ist es nun auch gelungen, IEEE als regelmäßigen Datenlieferant für *dblp* zu gewinnen. *dblp* erhält wöchentlich alle Neuaufnahmen der digitalen Bibliothek *IEEE Xplore*<sup>7</sup> und hat zudem auch vollen Zugriff auf deren Altbestand.

Diese Partnerschaften erwiesen sich im ersten Projektjahr für *dblp* als äußerst vorteilhaft und trugen erheblich zur Steigerung von Produktivität und Aktualität der Neuaufnahmen bei (vgl. Abschnitte 3.3 und 3.4). Es ist beabsichtigt, in Zukunft Kontakte mit weiteren Großverlagen zu intensivieren. Aber auch für die Autoren der Verlage dürfte sich die verbesserte Indexierung als Vorteil erweisen: Nicht selten werden inzwischen Publikation in *dblp*

<sup>6</sup><http://www.springerlink.com/>

<sup>7</sup><http://ieeexplore.ieee.org/>

indexiert, noch bevor sie in den digitalen Bibliotheken der Verlage gelistet werden.

### 3.9 Aufbau einer bibliometrischen Infrastruktur

Ein wesentliches Ziel des Projektes ist die Bildung einer substantiellen Infrastruktur, die langfristig eine umfassende Datenbasis für bibliometrische Analysen im Fach Informatik bereit stellt. Im ersten Projektjahr wurden in enger Absprache mit dem *dblp* Advisory Board sowie dem Fakultätentag Informatik einige Vorarbeiten zur Erreichung dieses Ziels geleistet.

Als erster Schritt zur Erprobung der Erhebung bibliometrischer Rankings wurde Anfang 2012 eine Serie von Umfragen unter den Seminarteilnehmern von Schloss Dagstuhl initiiert. Mittels der Umfragen konnten wertvolle Erfahrungen gesammelt werden. Vorbereitungen zur Durchführung einer Befragung eines großen nationalen oder internationalen Expertenkreises wurden eingeleitet.

Die konkrete Etablierung einer eigenen bibliometrischen Infrastruktur wird einen Schwerpunkt des zweiten Projektjahres darstellen.

### 3.10 Nutzerorientierte Neugestaltung der Webseiten

Die *dblp*-Webseiten wurden 1993 als einfache Sammlung statischer HTML-Seiten geschaffen. Seitdem hat sich die Erscheinung des Webfrontends von *dblp* in Prinzip nur wenig gewandelt. Das rustikale Tabellenlayout wirkt heute veraltet und fällt im optischen Vergleich und in Sachen Benutzerführung gegenüber modernen Websystemen wie Google Scholar oder Microsoft Academic Search zurück.

Aus diesem Grund wurde das Webfrontend überarbeitet und im Sinne eines leichtgewichtigen, modernen Designs unter Beibehaltung der inhaltliche Aussagekraft erneuert. Neben einer optischen Generalüberholung stand dabei insbesondere auch die nutzergerechte Aufbereitung der bibliographischen Daten im Vordergrund. So wird einer häufig geäußerten Kritik Rechnung getragen und neben der bisherigen chronologischen Sortierung der Autorensseiten auch eine kategorisierte Sortierung nach Publikationstyp angeboten. Des Weiteren ist bei der Restrukturierung auf einen modularen Aufbau geachtet worden, der in Zukunft die Erweiterung um zusätzliche Komponenten und Verlinkungen, sowie die Anpassung an unterschiedlichste Ausgabegeräte (z.B. Drucker oder Mobilgeräte) erleichtern soll.

Die Webseiten befinden sich derzeit in der Finalisierungsphase und sollen spätestens Anfang des dritten Quartals 2012 in den öffentlichen Testbetrieb übergehen.

## 4 Fazit

Die Zusammenarbeit von Schloss Dagstuhl mit der Literaturdatenbank *dblp* hat bereits nach Ende des ersten Projektjahres einige substantielle Erfolge vorzuweisen. Zudem findet die neue Zusammenarbeit auch in der internationalen Wissenschaftsgemeinschaft anklang. In Gesprächen mit Seminarteilnehmern auf Schloss Dagstuhl wird uns immer wieder positiver Zuspruch zugetragen.

Die durch Personalstärke und technische Weiterentwicklung ermöglichte Produktivitätssteigerung hat schon jetzt die ursprünglichen Erwartungen übertroffen. Ziel des zweiten Projektjahres ist, dieses Niveau zu halten.

Darüber hinaus wurde die organisatorische Infrastruktur von *dblp* auf ein neues Fundament gestellt. Das neu gegründete *dblp* Advisory Board hat die wissenschaftliche Aufsicht über *dblp* übernommen. Die für die Datenaufnahme nötige Priorisierung von Publikationsreihen hat durch die Einbeziehung von Experten und externen Rankings eine qualitative Verbesserung erfahren.

Für den Aufbau einer eigenen bibliometrischen Infrastruktur wurden erste Schritte unternommen. Die konkrete Etablierung einer Infrastruktur wird einen Schwerpunkt des zweiten Projektjahres darstellen.

Es ist zudem hervorzuheben, dass die Förderung durch die Klaus Tschira Stiftung ihren signifikanten Anteil an dem steten Voranschreiten des Projekts hat.

## Literatur

- [1] Oliver Hoffmann. Regelbasierte Extraktion und asymmetrische Fusion bibliographischer Informationen. Diplomarbeit, Universität Trier, 2009.
- [2] Oliver Hoffmann, Marc Herbstritt, and Christian Lindig. LZI+DBLP, Zwischenbericht, Mai 2011.