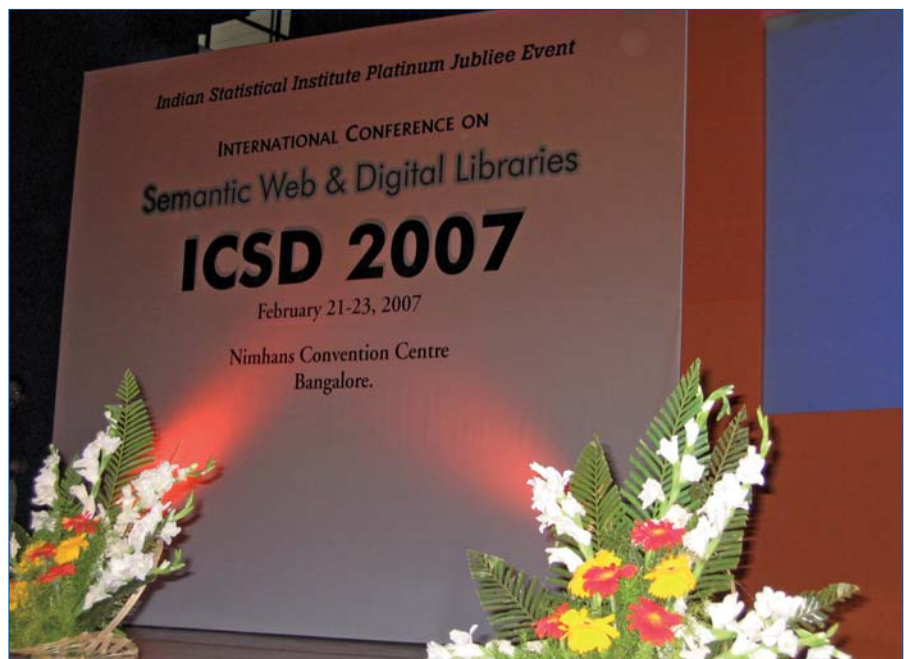


# International Conference on Semantic Web and Digital Libraries (ICSD) 2007 in Bangalore, Indien

von Rafael Ball

Das Indian Statistical Institute (ISI) in Bangalore, veranstaltete vom 21. bis 23. Februar 2007 die „International Conference on Semantic Web and Digital Libraries (ICSD)“. Mehr als dreihundert Teilnehmer aus zwölf Ländern waren zusammengekommen, um sich den Themen Semantic Web, Digital Libraries und den entsprechenden Schnittstellenthemen zu widmen. Der Veranstalter ist in der informationswissenschaftlichen Community nicht unbekannt: Das Indian Statistical Institute (ISI) wurde 1931 von Professor Shandra Maha Lanubis als Statistisches Institut zur Unterstützung gesamtstaatlicher Planungen in Indien gegründet. In den 60er Jahren kamen dann die Informations- und Dokumentationswissenschaften mit Lehre und Forschung sowie die besonderen Schwerpunkte Bibliothekswissenschaft und Bibliometrie hinzu. Der unvergessene und in Bangalore am ISI besonders hoch verehrte Bibliothekswissenschaftler Shiyali Ramamrita Ranganathan etablierte die bibliothekswissenschaftliche Forschung am Institut, das durch dessen Studien und Beiträge internationale Bedeutung erlangte. Das ISI ist heute aktiv in der Erforschung und Verbreitung von statistischem Wissen, der Entwicklung von statistischen Theorien und Methoden und der Unterstützung gesamtindischer Planungsaktivitäten. Zudem engagiert sich das ISI in den Natur- und Sozialwissenschaften im Bereich der Wissenschaftsevaluation und Bibliometrie. Ein dritter wichtiger Schwerpunkt sind heute die Themen Digital Library und Semantic Web, wie sie auch als Konferenzthema formuliert worden sind.

Das Indian Statistical Institute in Bangalore befindet sich in jener Stadt, die gemeinhin als „Silicone Valley“ von Indien bezeichnet wird. Mit fast fünf Millionen Einwohnern ist die Stadt heute ein wichtiges Zentrum der Luft- und Raumfahrtindustrie und seit einigen Jahren auch eines



Das „Bühnenbild“ bei der Konferenzeröffnung

der wichtigsten IT-Zentren des Landes, in dem sich viele in- und ausländische Computer- und Hochtechnologiefirmen in großen „Technikparks“ wie Electronics City oder International Technology Park angesiedelt haben. Entsprechend hoch ist das durchschnittliche Qualifikationsniveau in der Stadt und ebenso der Lebensstandard. Dennoch darf die für Indien außergewöhnlich breite Mittelschicht nicht darüber hinwegtäuschen, dass das Stadtzentrum und seine Infrastruktur ganz offensichtlich nicht den Anforderungen dieser hoch qualifizierten und überdurchschnittlich verdienenden indischen Informatiker gewachsen ist und auch in Bangalore noch immer fast 10% der Bevölkerung unterhalb des Existenzniveaus in Slums leben. Das Department of Documentation and Research Training Centre des ISI richtete vom 19. bis 21. Februar 2007 unter der Leitung von Professor Dr. A.R. D. Prasad sowie seiner Mitarbeiterin Dr. Devika P. Madalli die Konferenz ICSD aus (Bild 2).

Die Themen der Konferenz gruppierten sich dann auch um folgende Inhalte:

- Semantic Web
- E-Prints
- E-Governance
- Meta Data and Knowledge-Organisation
- Multilingual Digital Libraries
- Ontologies
- E-storage and Preservation
- Digital Library Practise
- Open Access
- Institutional Repositories Content
- Knowledge Management

Nach der Eröffnung der Konferenz (Bild 3), die mit einem in Indien weit verbreiteten, traditionellen Ritual des Anzündens von Kerzen und der Bitte um Erleuchtung für die Sprecher und Zuhörer begann, referierten Keynote-Speaker in ausführlichen Seminaren etwa zu den Themen Semantic Web oder E-Prints. Diese Seminare dauerten mehrere Stunden und ließen einen

vertieften Einblick in die Thematik zu, die Diskussion mit den Zuhörern entspannt sich workshopartig, so dass eine tiefe Durchdringung des Themas möglich war. Leider kamen dem gegenüber viele internationale Originalbeiträge nur mit sehr kurzen Redezeiten zu Zuge. Trotz allem, die hohe Zahl der Teilnehmer und die hohe Zahl der zugelassenen Vorträge in mehreren parallelen Veranstaltungssessions ließen eine andere Form der Präsentation innerhalb der drei Tage dauernden Konferenz leider nicht zu.

Die Konferenz stand im Spannungsfeld der Zentralthemen Semantic Web und Digital Libraries. So waren auch die bereits erwähnten Seminare diesen Themen gewidmet. Ivan Herman, ein ungarischer Mathematiker, der z. Zt. als Professor am W3C-Institut in Amsterdam arbeitet, berichtete ausführlich über die Möglichkeiten des Semantic Web und trat dabei mit den Teilnehmern auch in eine Diskussion ein über die Frage „was denn nun das Semantic Web eigentlich ausmache?“. Dabei wurde immer wieder klar, dass das Semantic Web, was nichts anderes bedeutet als eine Softwareanwendung zur Erschließung von Online-Inhalten und damit eigentlich nur Meta-Daten und Zusatzinformationen für intelligent gemachte Webinhalte präsentiert, selbstverständlich einen ganz eindeutigen Bezug zu digitalen Bibliotheken hat. Denn haben auch Bibliothekare in den letzten Jahrzehnten hervorragende Erschließungsinstrumente für ihre traditionellen gedruckten Bestände entwickelt, so fehlen diese Erschließungsinstrumente in der notwendigen Vollkommenheit nun für die digitalen Inhalte. Das Semantic Web könnte dazu neue Möglichkeiten bieten und so führen Tools der Kommunikations- und Interaktionssoftware des Semantic Web bereits zu den ersten Ergebnissen wie „social tagging“ „folksonomies“ und anderen.

Obwohl die Techniken des Semantic Web bereits seit Jahren existieren, werden sie erst heute von den Bibliothekaren zur Anwendung in digitale Bibliotheksbestände gebracht. Der Grund hierfür, so war die Meinung der IT-Fachleute und der Bibliothekare im Seminar von Ivan Herman, liegt darin, dass diese Erschließungstools des Semantic Web viel zu früh und noch zu wenig entwickelt in die Hände von Anwendern (Bibliothekaren) gelangt waren. Damit wurden sie schnell als unausgereift und zu



Blick in das Gästehaus des Indian Institute of Bangalore

komplex zur Seite gelegt. Sie waren von der IT-Seite her noch nicht mit der notwendigen Einfachheit programmiert worden. In einem zweiten Anlauf erhofft sich die Anwendergruppe der Digital Librarians eine nun erfolgreichere und fruchtbarere Zusammenarbeit mit den Softwareentwicklern im Bereich des Semantic Web. Allein schon für diese Diskussionen und für diesen Input hätte sich der Besuch der Konferenz gelohnt.

In einem zweiten, sehr maßgeblichen und umfangreichen Seminar, berichtete Leslie Carr (Informatiker an der University of Southampton und Mitarbeiter im EPrints Repository software Team) über die Möglichkeiten und den Einsatz von E-Prints in Großbritannien und darüber hinaus. Dies wäre an sich noch kein revolutionäres Thema, denn die Zurverfügungstellung von Publikationsplattformen in Zeiten von E-Science ist nichts mehr Neues. In Indien jedoch bedeutete die ausführliche Darstellung von E-Prints eine kleine Revolution.

Durch das unermüdliche Bestreben von Professor Prasad, dem Veranstalter der ICDS und seiner Zusammenarbeit mit dem Massachusetts Institute of Technology (MIT) bei der Entwicklung der Institutional Repository Software DSpace, wird nahezu flächendeckend über Gesamtindien die Software des MIT vertreten und eingesetzt. Es gibt kaum eine Universität in der nicht ein Dokumentenserver auf DSpace-Basis zu finden wäre. Umso spannender war das Seminar von Leslie Carr, der die Vorzüge und Nutzungsmöglichkeiten von E-Prints in der gebotenen sachlichen Tiefe und gleichzeitigen Allgemeinverständlichkeit vorführte. Da in Indien bereits mehr als eine Million Bücher digital vorliegen, sind diese Themen immer „hot topics“ für die Bibliothekarinnen und Bibliothekare.

Überraschende Einhelligkeit herrschte unter allen Keynote-Speakern und Seminarleitern bei der Auffassung, dass das Web 2.0 eigentlich nur eine Vision, ein neuer Hype ohne eigentlichen Inhalt sei. So

sprach Ivan Herman davon, dass das Web 2.0 nichts anderes sei als die Sammlung von neueren Kommunikations- und Interaktionssoftwaretools, während Professor Wendy Hall, Department of Electronics & Computer Science, University of Southampton, UK als Keynote-Speakerin die Vermutung äußerte „The Semantic Web is still a dream!“. Gleichzeitig jedoch hatte sie Web 2.0 bereits abgehakt und diskutierte über die Möglichkeiten von Web 3.0, dessen Konturen jedoch noch weitaus unschärfer seien als diejenigen des Web 2.0.



Open Access ist überall ein Thema

Beispielhaft für die verschiedenen thematischen Sessions sollen jeweils ein bis zwei Vorträge referiert werden. So sprach Frederic J. Friend (University College London) in der Session „Knowledge Organisation and Information Retrieval“ über die Bedeutung des bibliothekarischen Grundwerkes von S. R. Ranganathan „The Future of Libraries in the World“. Ranganathan, der als Begründer der Bibliothekswissenschaft und der Dokumentation am Indian Statistical Institute in Bangalore gilt und nicht nur in Indien eine bibliothekarische Institution darstellt, hat in seinem Grundlagenwerk „The Five Laws of Library Science“ 1931 zentrale Grundstrukturen des modernen Bibliotheksmanagements gelegt. Friend konnte in seinem Vortrag nachweisen und sehr eindringlich aufzeigen, dass die fünf Gesetze von Ranganathan auch in Zeiten der Digital Library Gültigkeit haben. Das erste Gesetz „Books are for the use“ ist auch heute noch in Zeiten digitaler Information von großer Bedeutung. Die digitalen Informationen müssen genutzt werden können, die technischen, administrativen und rechtlichen Barrieren geschleift und der Zugang zu Informationen für alle möglich gemacht werden. Die vor mehr als 70 Jahren formulierte zweite Regel „Every person his or her book“ interpretierte Friend in Zeiten der digitalen Bibliothek als die Notwendigkeit, dass Bibliothekarinnen und Bibliothekare mit dazu beitragen, den digital divide zu überwinden, die bislang keinen oder nur erschwerten Zugang zu digitalen Informationen haben, etwa durch solarbetriebene Laptops, die Zurverfügungstellung von Sammelpunkten mit Internetanschluss und ähnlichen Projekten.

Auch das dritte Gesetz „Every book it's reader“ hat in Zeiten der digitalen Informationsschwemme größte Aktualität. Es darf und muss heute durchaus gefragt werden, ob der immense Inhalt, der elektronisch



Der Veranstaltungsort der ICSD: Das Convention Center in Bangalore

zur Verfügung steht, überhaupt in angemessener Weise rezipiert werden kann. Während das vierte Gesetz von Ranganathan („save the time of the reader“) angesichts immer größer werdender Softwarepakete und entsprechender Ladezeiten im digitalen Zeitalter traurige Aktualität erlangt, meinte es 1931 noch die Verkürzung der Wartezeiten für die Besorgung von Literatur aus dem Magazin oder entfernten Bibliotheken. Heute aber ist eher die Realisierung einer benutzerfreundlichen Oberfläche, die Einheitlichkeit von Formaten und die gute Lesbarkeit von digitalen Inhalten mit dieser Formel gemeint. Auch das letzte Gesetz von Ranganathan („a library is a growing organism“) ist heute in Zeiten der Kooperation digitaler Bibliotheken regional, national und international von höchster Aktualität. Die Tatsache, dass die fünf Grundgesetze von Ranganathan auch 70 Jahre nach der Veröffentlichung und nach der digitalen Revolution noch immer Gültigkeit haben, so sieht es Friend, beweise den Genius in der Arbeit von S. R. Ranganathan.

In der Session „Digital Libraries and Institutional Repositories“ soll auf den Vortrag

von T. D. Webb und Bin Zhang „The role of scholarly publishing in digital libraries“ eingegangen werden. Webb, Bibliothekar an der California State University in Sacramento und Zhang, Bibliothekar an der renommierten Hongkong Baptist University, berichten dabei über die Möglichkeiten der wissenschaftlichen Kommunikation in Zeiten von Digital Libraries und E-Publishing. Dabei unterscheiden die Autoren zwischen verschiedenen digitalen wissenschaftlichen Inhalten:

- **born again** sind Informationen und Quellen, die als gedruckte Ressourcen digitalisiert und zur Verfügung gestellt werden.
- **born digital** Informationen waren niemals als gedruckte Quellen verfügbar waren und existieren nur in elektronischer Form.

Die University of California at Berkley beispielsweise unterscheidet hierbei vier Ebenen der Langzeitarchivierung (Bild 5). Die erste Ebene sind archivierte, elektronische Informationen. Das sind solche, die fest auf den Servern der Universität installiert sind und dort auch unbegrenzt vorgehalten werden. Danach gibt es sogenannte „bediente Informationen (served informations)“. Das sind solche, die nur vorübergehend auf den Servern der Universität vorgehalten werden, während „mirrored informations“, also gespiegelte Daten, solche sind, die von einer anderen Quelle zur Verfügung gestellt und gespiegelt werden. „Linked information“, also solche zu denen ein Link hergestellt wird, sind Daten auf der untersten elektronischen Ebene im digitalen Information Management der University of California.

Im Resümee ziehen die beiden Autoren den Schluss, dass Bibliotheken de facto digitale Verleger geworden und gleichzeitig einige Verleger zu Informationsanbietern wie Google, Yahoo oder Open Content Alliance geworden seien.

Social Event: Eine indische Oper zum Abschluss der Eröffnung





Digital Preservation: Ein zentrales Thema auf der ICSD in Bangalore



Ivan Herman vom W3C-Konsortium bei seinem Tutorial

Das große Land Indien hat neben den technischen Errungenschaften im Bereich der digitalen Bibliotheken eine höchst komplexe Struktur der Nationalsprachen. In Indien gibt es mehr als 428 Sprachen mit 1.600 Dialekten und entsprechenden Schriftsystemen. Für die Aufbereitung und die Suche von digitalen Informationen in digitalen Bibliotheken ist die Sprachenvielfalt ein echtes Problem. Insbesondere deshalb, weil sich die komplexen Schriftzeichen der meisten indischen Sprachen nicht einfach abbilden lassen. Professor Aditya Tripathi und Sneha Tripathi von der Banaras Hindu University in Varanasi berichteten über das Projekt einer multilingualen digitalen Bibliothek mit verschiedenen indischen Sprachen. Man beginnt erst jetzt für die meisten indischen Sprachen OCR-Programme zu entwickeln. Unter dem großen Themenblock „Open Access“ und „Institutional Repositories“ berichteten Swapan Deoghuria und Satyabrata Roy von der Indian Association for Cultivation of Science Jadavpur in Kolkata über ihre Umfrage zu Open Access (OA) in einer Wissenschaftscommunity. Mehr als 300 Fragebögen wurden an Wissenschaftlerinnen und Wissenschaftler der Universität verschickt, um die Akzeptanz, den Einsatz und die Bewertung von Open Access durch die Wissenschaftler zu erfragen. Ganz ähnlich wie in europäischen Studien zeigte sich auch hier, dass ein großer Anteil von Wissenschaftlern (80 %) Open Access nur nutzt, um Literatur zu lesen und downzuladen, während weniger als 20 % sowohl publizieren als auch Literatur nutzen. Rund ein Drittel aller Wissenschaftler drückte die Bereitschaft aus, gelegentlich in Open Access zu publizieren und nur ein verschwindend geringer Bruchteil von Wissenschaftlern kannte Open Access nicht oder sah die Entwicklung kritisch. Vor dem Hintergrund von Wissenschaftskommunikation hielten rund 50 % der Wissenschaftler OA für geeigneter, wenn es um die schnelle Verbrei-



Bangalore – High Tech und überfüllte Straßen

tung ihrer Ergebnisse ging, 30 % der Wissenschaftler glaubten, mit OA eine breitere Rezipientenbasis zu erreichen als mit traditionellem Publikationsverhalten. Nahezu alle Wissenschaftler aber (und dies steht im Gegensatz zu den Publikationen von Harnard und Eisenbach, die behaupten, dass Open Access eine höhere Wahrnehmung erzielt als traditionelle Publikationen) sind der Meinung, dass die traditionellen Veröffentlichungen eine höhere Wahrnehmung erzielen als eine Open Access-Publikation.

Die Konferenz brachte auf gelungene Weise Spezialisten aus der IT-Branche und dem Bibliothekswesen mit den zentralen Themen des 21. Jahrhunderts „Semantic Web“ und „Digital Libraries“ zusammen. Dankenswerterweise haben die Veranstalter des Indian Statistical Institute in Bangalore den Teilnehmern bereits zur Konferenz einen gut aufbereiteten Proceedingsband inklusive einer CD-ROM mit auf den

Weg gegeben (International Conference on Semantic Web & Digital Libraries (ICSD 2007), 21 – 23 February 2007, A.R.D. Prasad; Devika P. Madalli, Bangalore, India, Documentation Research & Training Centre, Indian Statistical Institute, Indian Statistical Institute Plantinum Jubilee Conference Series), in dem die vielen guten Beiträge nachgelesen werden können.

#### Zum Autor:

**Dr. Rafael Ball**  
ist Leiter der  
Zentralbibliothek des  
Forschungszentrums  
Jülich GmbH  
D-52425 Jülich  
E-Mail:  
[r.ball@fz-juelich.de](mailto:r.ball@fz-juelich.de)

