

# io-port.net - das Informationsportal für die Informatik

VON DIPL.-KFFR., DIPL.-WIRT.-ING. (FH)  
REBECCA BULANDER,  
DIPL.-KFFR. AGNES KOSCHMIDER,  
PROF. DR. ANDREAS OBERWEIS<sup>1</sup>,  
DIPL.-BIOL. UTE RUSNAK<sup>2</sup>

<sup>1</sup> Institut für Angewandte Informatik und  
Formale Beschreibungsverfahren (AIFB),  
Universität Karlsruhe (TH)

<sup>2</sup> FIZ Karlsruhe - Fachinformationszentrum  
Karlsruhe

**D**as Informatikportal io-port.net wird seit Januar 2006 online angeboten. Es wurde im Rahmen eines vom Bundesministerium für Bildung und Forschung (BMBF) geförderten Projekts als Informationsportal für die Informatik entwickelt. io-port.net bietet über eine zentrale Web-Oberfläche einen komfortablen Zugriff auf alle relevanten Fachpublikationen der Informatik. Das Nachweissystem für Publikationen umfasst derzeit mehr als zwei Millionen bibliographische Einträge. Darüber hinaus werden Nutzer von io-port.net durch semantische Werkzeuge bei einer effizienten Erschließung von Literaturnachweisen unterstützt. Der folgende Beitrag beschreibt die wesentlichen Funktionalitäten und Komponenten des Informatikportals io-port.net. Außerdem werden Alleinstellungsmerkmale von io-port.net gegenüber anderen Suchdiensten aufgezeigt.

## Überblick über io-port.net

An dem BMBF-geförderten Projekt Fachinformationssystem Informatik (FIS-I) waren neben der Gesellschaft für Informatik e. V. (GI) und dem Fachinformationszentrum Karlsruhe, Wissenschaftler der Universitäten Karlsruhe und Trier sowie der Technischen Universität München beteiligt. Das Informationsportal io-port.net ermöglicht den Zugang zu einem umfassenden Kompetenz- und Dienstleistungsnetz für die Informatik und deren verwandten Gebieten [1]. Das Kernstück von io-port.net bildet ein Nachweissystem für bibliographische Publikationen, das auf den bereits existierenden Beständen der Nachweisdatenbanken CompuScience, DBLP, LEABiB und The Collection of Computer Science Bibliographies (CCSB) aufbaut. Dabei war eine der Herausforderungen im FIS-I-Projekt, die verschiedenen Metadaten-Beschreibungsformate der Projektpartner zu einem Austauschformat zu vereinheitlichen.

Im Moment weist io-port.net mehr als zwei Millionen bibliographische Einträge strukturiert und standardisiert nach. Ausgehend vom Nachweissystem wird dem Nutzer von io-port.net direkt im Anschluss an eine Recherche eine Vermittlung der jeweiligen Volltexte angeboten, die er dann auch lokal auf seinem Rechner speichern kann. Besonders hervorzuheben ist der Zugriff auf elektronische Volltexte der gesamten LNI-Reihe (Lecture Notes in Informatics) der Gesellschaft für Informatik, in der zahlreiche Tagungsbände veröffentlicht werden, sowie

weitere elektronische Beiträge aus GI-Zeitschriften. Geplant ist auch eine Volltextvermittlung zu Diplomarbeiten und Dissertationen aus der Informatik.

Zusätzlich zu diesem Hauptkomplex der Literatursuche und Volltextvermittlung sind weitere Fachinformationen unterschiedlicher Art in das Portal integriert. So ermöglicht z. B. der io-port.net-Editor die Erfassung und die Pflege persönlicher Publikationslisten. Weiterhin wurden in io-port.net semantische Werkzeuge integriert, die in dem BMBF-Partnerprojekt Semantische Methoden und Werkzeuge für Informationsportale (SemlPort) entwickelt wurden [2].

So machen die Projektergebnisse von FIS-I und SemlPort io-port.net zu einem nutzerorientierten und umfassenden Informationsportal für die Informatik.

io-port.net bietet Nutzern für die Informationsdarstellung kostenfreie Basis- und kostenpflichtige Mehrwertdienste an. Innerhalb der Basisdienste kann der Nutzer einfache Suchen über den gesamten bibliographischen Bestand durchführen; ebenso hat er Zugriff auf das Informatiklexikon der GI, auf diverse Linksammlungen zu speziellen Bereichen der Informatik sowie auf Publikationslisten, die von anderen io-port.net-Nutzern veröffentlicht wurden. Für ausgewählte Fachbereiche der Informatik werden darüber hinaus redaktionell betreute Themenseiten geführt, die einen kompakten Überblick über aktuelle Forschungs- und Entwicklungsergebnisse, laufende Projekte, Publikationen sowie aktuelle Veranstaltungen geben. Erweiterte Recherchemöglich-



▲ **Abbildung 1: Das Informatikportal io-port.net**

keiten (Mehrfeldsuche) sowie personalisierte Dienste, wie zum Beispiel Push-Dienste oder die Erstellung eigener Publikationslisten erhalten Nutzer im Bereich der Mehrwertdienste; hier werden Nutzer außerdem bei ihrer Recherche durch semantische Werkzeuge unterstützt (siehe Absatz: Integration semantischer Werkzeuge in io-port.net). Abb. 1 zeigt die Einstiegsseite von io-port.net.

#### Suchfunktionalitäten in io-port.net

Den Benutzern von io-port.net stehen bei ihrer Recherche so genannte Einfach- und Mehrfeldsuchen zur Verfügung. Damit sind sowohl einfache als auch komplexe Anfragen an die Datenbasis möglich wie z. B. Verwendung verschiedener Suchkriterien und logischer Verknüpfungen, Phrasensuche und Trunkierung. Als Trunkierung wird die Abkürzung von Suchbegriffen bei der

Recherche in Datenbanken und Information-Retrieval-Systemen bezeichnet. Nach der Anzeige der Suchergebnisse können Nutzer sich die Publikationen über einen automatischen Vermittlungsdienst bestellen oder sich direkt zu den entsprechenden Volltexten vermitteln lassen. In Abbildung 2 wird die Mehrfeldsuche dargestellt.

#### Architektur von io-port.net

Die Architektur des Gesamtsystems sieht eine Zweiteilung in ein Daten-Managementsystem und ein Portalsystem vor (siehe Abb. 3). Im Daten-Managementsystem werden die Daten gesammelt, aufbereitet und archiviert. Über Komponenten des Portalsystems werden Nutzern Daten mit verschiedenen Abfragemechanismen angeboten. Das io-port.net-Portalsystem stellt eine webbasierte Oberfläche zum Suchen und Navigieren nach statischen und dynamischen

Inhalten bereit. Der Zugang zu Basisdiensten wie der Einfachsuche erfordert keine Authentifizierung des Benutzers - im Gegensatz zu Mehrwertdiensten wie der Mehrfeldsuche. Die vom Nutzer formulierten Anfragen werden über eine generische Anfragekomponente je nach Anwendungsfall an spezialisierte Anfragekomponenten weitergeleitet, die auf lokale Datenbestände zugreifen. Die Personalisierungskomponente unterstützt Nutzer durch das Angebot eines Push-Dienstes, individuelle Suchanfragen zu Interessensgebieten zu speichern und zu pflegen. Über Push-Technologien (ähnlich einer Benachrichtigungsfunktion) werden die Nutzer automatisch über entsprechende neu aufgenommene Publikationen informiert.

In dem Kundenverwaltungs- und Abrechnungssystem, welches mit einem Enterprise Resource Planning-System (ERP-System) verbunden ist, wird die Verrechnung



▲ **Abbildung 2: Mehrfeldsuche in io-port.net**

kostenpflichtiger Mehrwertdienste durchgeführt.

Für die Betreuung von Themenseiten durch Redakteure wurde in das Portalsystem ein eigenes Content Management System (CMS) integriert. Weitere externe (nicht lokale) Datenquellen werden mittels eines QueryBrokers, einer speziellen Anfragekomponente für verteilte Datenquellen, in das Portal eingebunden.

Das Daten-Managementsystem übernimmt die Sammlung, Aufbereitung und Pflege der Daten für die zentrale Nachweisdatenbank. Spezielle Komponenten übernehmen die Kontrolle und Verbesserung der Datenqualität, wie z. B. die Vereinheitlichung bei Namensgebungen und die Dublettenerkennung und deren Verarbeitung. Für den Austausch der bibliographischen Metadaten zwischen den Datenlieferanten und der zentralen Nachweisdatenbank wurde das gemeinsame Austauschformat Data Exchange Format 2 (DXF2) vereinbart. Über die DXF2-Komponente stellen die Projektpartner ihre Daten für den Import in io-port.net zur Verfügung. Neben der traditionellen manuellen Datenerfassung wird

der Bestand der io-port.net-Nachweisdatenbank auch durch einen ontologiebasierten Crawler, der im Rahmen des SemIPort-Projekts entwickelt wurde, kontinuierlich erweitert.

#### **Integriertes Nachweissystem**

io-port.net verwendet eine tiefe Datenintegration, die über eine für viele Portale typische lose Kopplung verschiedener Anbieter weit hinausgeht, und ermöglicht somit Nutzern einheitliche Anfragen über den Datenbestand der vier Partnerdatenbanken in definierter Qualität. Die Daten werden in einer zentralen Nachweisdatenbank zusammengeführt, wobei die Herkunft der Daten sichtbar bleibt. Jeder Datenlieferant verwendet bei der Strukturierung seiner Daten - der beschreibenden Metadaten - intern unterschiedliche Dateiformate, weshalb DXF2 als ein einheitliches Austauschformat für die Integration der Daten vereinbart wurde. Im Verlauf des Projektes wurden verschiedene Verfahrensweisen und Konvertierungsprogramme entwickelt, um die Daten in das DXF2-Format zu übertragen.

Eine große Herausforderung stellten die unterschiedlichen Quellenangaben ("sources") von Dokumenten dar. Bei älteren Dokumenten waren oft mehrere Angaben in einem Feld zusammengefasst. Dieses Feld sah z. B. für einen Zeitschriftenartikel folgendermaßen aus:

*"SO: Theor. Comput. Sci 274, No.1-2, 3-41 (2002). [ISSN 0304-3975]"*

Das DXF2-Format verlangt jedoch eine Differenzierung nach einzelnen Feldern, in diesem Fall: Journal, Volume, Number, Page, Year und ISSN. Zur Trennung dieser Felder wurden AWK-Programme erstellt, welche für die Bearbeitung von Textdateien geeignet sind. Teilweise gestaltete sich diese Trennung bei Büchern oder Sammelbänden noch komplizierter. Zusätzlich waren Nachbesserungen von Hand notwendig.

#### **Integration semantischer Werkzeuge in io-port.net**

Eine Studie zum Thema "Nutzung elektronischer wissenschaftlicher Informationen in der Hochschulausbildung" hat auf die durchschnittliche bis schlechte Informa-

tionskompetenz von Studierenden hingewiesen [3]. Die Befragten gaben an, dass ihre freien Suchen mittels Suchmaschinen zu keinem zufrieden stellenden Ergebnis geführt hätten; dies sei teilweise auf eine unsystematische und laienhafte Suche zurückzuführen. Aus diesem Grund lag ein Hauptaugenmerk im Projekt FIS-I in der Integration semantischer Werkzeuge zur Erschließung von Literaturnachweisen sowie zur Gestaltung intelligenter Nutzerschnittstellen (siehe Abb. 4). Semantische Werkzeuge sollen Nutzer unterstützen, den Suchprozess zu verkürzen und dabei qualitativ hochwertigere Suchergebnisse zu erzielen. Hierfür bietet io-port.net dem Nutzer verschiedene Funktionskomponenten wie Digital Bibliographic Library (DBL)-Browser [4], Recommender- und Document-Manager-Systeme [5] sowie Query Refinement [6] an.

Der DBL-Browser ist ein intuitiv nutzba-

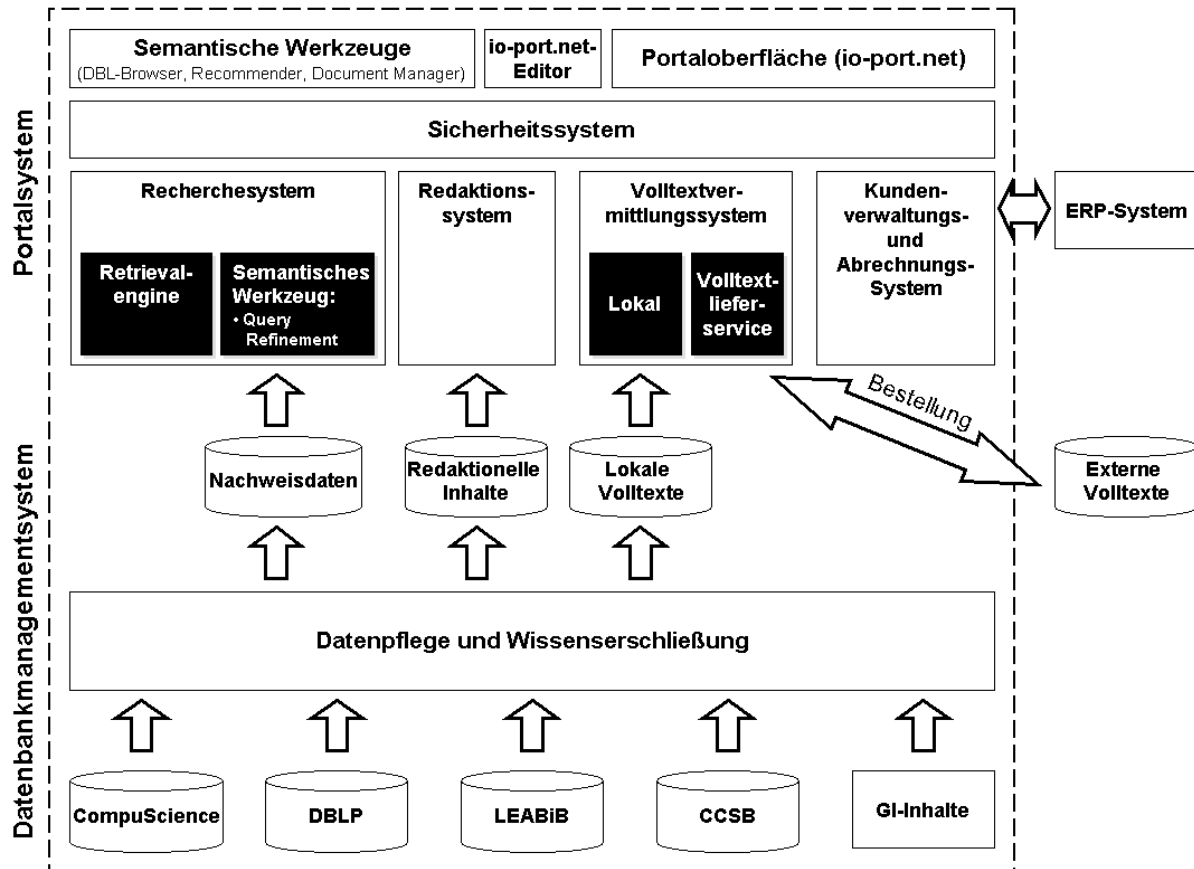
res grafisches Werkzeug zum Navigieren in der io-port.net-Datenbank; er kann lokal auf dem Rechner installiert werden. Der Document-Manager ermöglicht dem Nutzer, lokal gespeicherte Dokumente oder URLs zu verwalten, zu kommentieren und Beziehungen zwischen den Dokumenten und URLs z. B. als Gruppierungen zu spezifizieren. Die modellierten Beziehungen zwischen den Dokumenten (Dokumentenkontext) können in io-port.net gespeichert und anschließend durch die Recommender-Komponente verwendet werden. Der Recommender sortiert die Suchergebnisse entsprechend dem vorher definierten Dokumentenkontext nach Relevanz für den Nutzer.

Ein weiteres semantisches Werkzeug ist Query Refinement. Mit Hilfe dieses Werkzeuges können Suchanfragen durch vorgeschlagene relevante Begriffe präzisiert und verfeinert werden. Die Suchanfrage nach

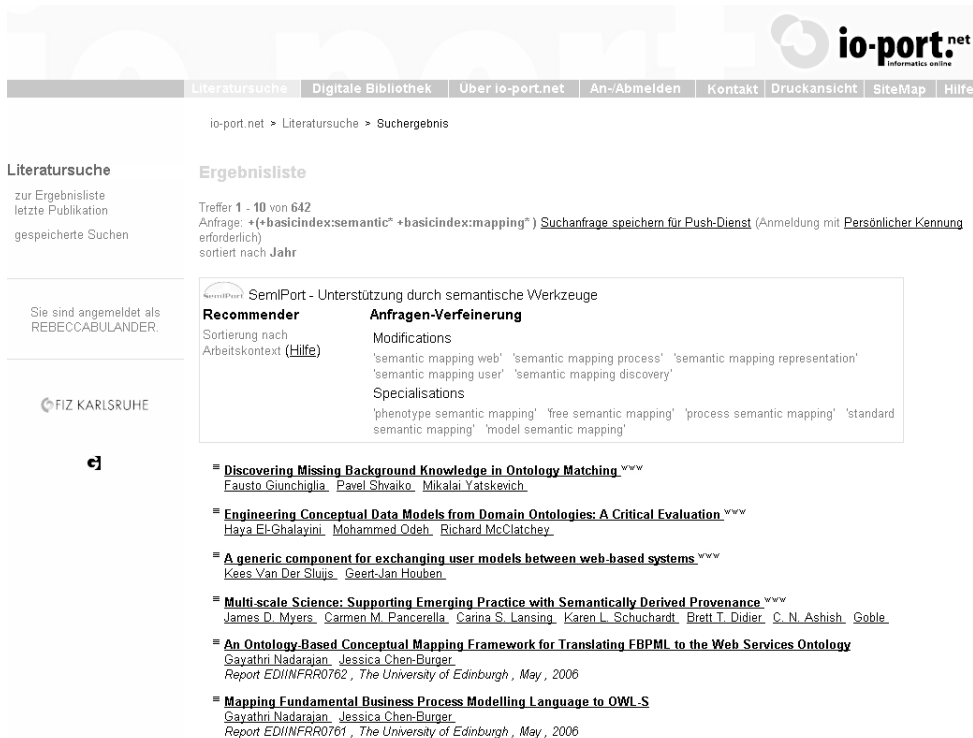
"Business Process" verfeinert das Query Refinement-Werkzeug unter anderem zu "Business Process Management" oder "Business Process Intelligence".

### Google versus io-port.net

Seit November 2004 bietet Google unter scholar.google.com einen speziellen Suchdienst für wissenschaftliche Recherchen an. Google Scholar durchsucht nach Angaben der Betreiber Bücher, wissenschaftliche Abhandlungen, technische Dokumente, Fachzeitschriften und sonstige Literatur aus dem Umfeld von Forschung und Lehre. Einen bedeutenden Vorteil gegenüber dem "klassischen" Google erlangt dieser Suchdienst durch die Verwendung einer Zitationsanalyse für Publikationen (Cited by) [7]. Allerdings ist der Suchalgorithmus noch nicht effizient, da Suchergebnisse nicht immer zu Volltexten führen.



▲ Abbildung 3: Schematische Darstellung der Architektur von io-port.net



▲ **Abbildung 4: Semantische Werkzeuge in io-port.net**

Das Informationsportal io-port.net hat sich mit seinen Informationsangeboten und integrierten Suchfunktionalitäten im Gegensatz zu Google Scholar auf ein Fachgebiet spezialisiert - nämlich die Informatik - und deckt dieses Gebiet umfassend ab. Die Literaturnachweise sind aufgrund vollständiger Angaben und Abstracts qualitativ hochwertig; nach der Anzeige von Suchergebnissen gelangen Nutzer direkt zur Volltextvermittlung. Ein weiterer Vorteil von io-port.net besteht in den Hinweisen zu rechtlichen Fragestellungen bei Verweisen auf Volltexte und Volltextdownload, so dass Nutzer bei ihrer Recherche informiert werden.

Sowohl Google Scholar als auch io-port.net befinden sich momentan in der Startphase und müssen sich gegenüber verschiedenen Herausforderungen positionieren (u. a. Dublettenerkennung, Integration von Volltexten usw.). Insgesamt kann festgehalten werden, dass Google Scholar die Recherche in einem hochwertigen Informationsportal mit umfassenden und zuverlässigen Suchergebnissen wie io-port.net auf keinen Fall ersetzen kann.

### Stand der Arbeit

Das vom BMBF geförderte Projekt zu Entwicklung und Aufbau von io-port.net wurde von September 2002 bis März 2006 durchgeführt. Das Informatikportal io-port.net wird seit Januar 2006 von FIZ Karlsruhe als Informationsdienst angeboten. Das nachhaltige Angebot des Informationsdienstes erfordert neben dem zuverlässigen Betrieb auch die kontinuierliche Aktualisierung der Inhalte, die Erfassung neuer Metadaten sowie die Weiterentwicklung durch neue Funktionalitäten wie beispielsweise eine Diplomarbeitsbörse. Die kostenpflichtige Nutzung der Mehrwertdienste erfolgt über Jahreslizenzen, die von Privatpersonen oder Unternehmen und Organisationen (z. B. Universitäten) erworben werden können. Lizenzpflichtige Volltexte werden über einen automatischen Volltextdienst angeboten. Institutionen, die über einen eigenen Volltextdienst mit Zugriff auf bereits lizenzierte Volltexte wie beispielsweise ACM, IEEE oder Springer verfügen, können diese über eine geeignete Schnittstelle in io-port.net einbinden und über das Informatiksystem nutzen. Weitere Informationen zum Informations-

portal sind unter [www.io-port.net](http://www.io-port.net) zu finden. ■

### LITERATUR

[1] **Koschmider, A. et al. (2004):** Entwicklung eines Informationsportals für die Informatik, In: Dadam, P. und Reichert, M., Informatik verbindet, GI-Jahrestagung, Ulm, Band 2 (208-213), September 2004.

[2] **Agarwal, S. et al. (2003):** Semantic Methods and Tools for Information Portals, In: Dittrich, K. R.; König, W.; Oberweis, A.; Rannenber K. und Wahlster, W. (Hrsg.), Innovative Informatikanwendungen, GI-Jahrestagung, Frankfurt, Band 1 (116- 131),

September 2003.

[3] **SteFi (2001):** Nutzung elektronischer wissenschaftlicher Informationen in der Hochschulausbildung, Abruf am 20.02.2006, URL: [www.stefi.de](http://www.stefi.de)

[4] **Klink, S. et al (2004):** Visualising and Mining Digital Bibliographic Data, In: Dadam P.; Reichert, M. (Hrsg.), INFORMATIK 2004 - Informatik verbindet, GI-Jahrestagung Band 2 (193-197), September 2004.

[5] **Schwarzkopf, E.:** On the Construction of a Probabilistic Model for Assessing Users' Beliefs in an Information Environment, In: Dadam P.; Reichert, M. (Hrsg.), INFORMATIK 2004 - Informatik verbindet, Band 2 (176-181), September 2004.

[6] **Stojanovic, N.:** On discovery user's needs in the ontology-based portals using implicit relevance feedback, In: Dadam P.; Reichert, M. (Hrsg.), INFORMATIK 2004 - Informatik verbindet, GI-Jahrestagung Band 2 (182-186), September 2004.

[7] **Cusomano, M.A.: Google (2005):** What it is and what it is not. Communication of the ACM 48(2): 15-17, 2005.