

Editorial

Andreas Henrich · Gerhard Heyer · Christoph Schlieder ·
Theo Härder

Online publiziert: 19. Februar 2015
© Springer-Verlag Berlin Heidelberg 2015

Schwerpunktthema: Informationsmanagement für Digital Humanities

Mit dem vorliegenden Themenheft wollen wir ein Feld beleuchten, dessen Wurzeln bis in die Anfänge der Informatik selbst zurück reichen. Die Nutzung IT-gestützter Verfahren in den Geisteswissenschaften ist keineswegs neu. Allerdings haben sich die sogenannten Digital Humanities in den letzten Jahren durch verschiedene Entwicklungen (Leistungssteigerungen bei Hard- und Software, Fortschritte in Methoden und Algorithmen und nicht zuletzt der Aufbau von Infrastrukturen durch entsprechende Förderprogramme) sowohl quantitativ als auch qualitativ weiterentwickelt. Für die Geisteswissenschaften bringt diese Entwicklung neue Möglichkeiten zur Bearbeitung ihrer Forschungsfragen mit sich. Parallel entstehen zum Teil neue und spezifische Anforderungen für die Informatik. Das Datenmanagement und die Analyse der Daten bilden dabei wichtige Bereiche der Digital Humanities, die in diesem Themenheft exemplarisch an acht Beispielen dargestellt werden sollen.

Der Begriff der Digital Humanities (bisweilen auch als eHumanities oder abkürzend DH bezeichnet) ist so facettenreich und schwer abzugrenzen wie der Begriff der Geisteswissenschaften selbst. Zum einen liegt dies an der Heterogenität der geisteswissenschaftlichen Gegenstände und der Vielschichtigkeit der eingesetzten Methoden, zum anderen sind aber auch unterschiedliche Betrachtungsrichtungen möglich:

Der Begriff der Digital Humanities (bisweilen auch als eHumanities oder abkürzend DH bezeichnet) ist so facettenreich und schwer abzugrenzen wie der Begriff der Geisteswissenschaften selbst. Zum einen liegt dies an der Heterogenität der geisteswissenschaftlichen Gegenstände und der Vielschichtigkeit der eingesetzten Methoden, zum anderen sind aber auch unterschiedliche Betrachtungsrichtungen möglich:

- Eine Betrachtungsweise (und zwar die, die in diesem Themenheft auch weitestgehend eingenommen wird) bezieht sich auf die Unterstützung geisteswissenschaftlicher Forschungsarbeiten mit Methoden der digitalen Informationsverarbeitung sowie Forschungsfragen, die sich dadurch auch für die Informatik ergeben. Hier sind Aspekte der automatischen Textanalyse ebenso zu nennen wie Ansätze zum Forschungsdatenmanagement, der Analyse von 3D-Modellen oder der Geoinformatik.
- In einer zweiten Betrachtungsweise wird die Informatik selbst bzw. die Software- und Systementwicklung Gegenstand geisteswissenschaftlicher Forschung. Dabei wird die Geschichte der Informatik und der historische Kontext ihrer Entstehung ebenso reflektiert wie digitale Kunst oder die Rolle von Texten (Dokumentation, Quellcode etc.) sowie Interaktionsstilen in der Informatik. Solche ohne Zweifel wichtigen Fragestellungen sind aber von der oben genannten Betrachtungsweise abzugrenzen.
- Schließlich stellt auch die soziologische, philosophische oder anthropologische Betrachtung der Auswirkungen einer zunehmenden Digitalisierung ein wichtiges Forschungsgebiet dar, das aber ebenfalls von einer Unterstützung geisteswissenschaftlicher Forschungsarbeit mit digitalen Methoden abzugrenzen ist. Dabei wird es – als

T. Härder (✉)
TU Kaiserslautern, 67663 Kaiserslautern
Deutschland
E-Mail: haerder@cs.uni-kl.de

A. Henrich
Fakultät WIAI, Medieninformatik, Otto-Friedrich-Universität
Bamberg, An der Weberei 5, 96047 Bamberg, Deutschland
E-Mail: andreas.henrich@uni-bamberg.de

G. Heyer
Institut für Informatik, Automatische Sprachverarbeitung, Universität
Leipzig, Augustusplatz 10, 04109 Leipzig, Deutschland
E-Mail: heyer@informatik.uni-leipzig.de

Ch. Schlieder
Fakultät WIAI, Angewandte Informatik in den Kultur-, Geschichts-
und Geowissenschaften, Otto-Friedrich-Universität Bamberg, An der
Weberei 5, 96047 Bamberg, Deutschland
E-Mail: christoph.schlieder@uni-bamberg.de

Begleitforschung zu den Digital Humanities in der ersten oben genannten Betrachtungsweise – sehr wichtig sein, in wissenschaftssoziologischer Betrachtungsweise die Auswirkungen der Digitalisierung auf die geisteswissenschaftliche Forschung zu betrachten.

Wenngleich wir im Hinblick auf die Zielsetzung des Datenbank-Spektrums in diesem Themenheft den ersten Blickwinkel einnehmen wollen, wäre eine Einstufung der Digital Humanities als „weitere Bindestrich-Informatik“ zu kurz gegriffen. Um die Betrachtung geisteswissenschaftlicher Forschungsfragen im Hinblick auf ihre Unterstützung durch digitale Methoden zielgerichtet durchführen zu können, ist eine methodische Kluft zu überwinden. Während bei einer Ingenieurinformatik und letztlich auch bei einer Wirtschaftsinformatik von der Methodik und von den betrachteten Forschungsfragen her eine gewisse Nähe zwischen der Informatik und dem jeweiligen Anwendungsbereich existiert, sind die Unterschiede zwischen der Informatik und „den Geisteswissenschaften“ an dieser Stelle deutlich größer.

Aus der Sicht der Datenbanktechnologie und des Information Retrieval stellt sich nun die Frage, welche neuen Herausforderungen sich aus den Digital Humanities für das Informationsmanagement ergeben.

Facetten der „digitalen“ Geisteswissenschaften

Der Begriff der Geisteswissenschaften selbst ist in der Abgrenzung Gegenstand verschiedenster Diskussionen. Im Verständnis der Digital Humanities reicht das Spektrum in der Regel über Geschichts-, Kunst-, Literatur-, Religions- und Sprachwissenschaften hinaus und umfasst auch kulturwissenschaftliche, humanwissenschaftliche und sozialwissenschaftliche Fächer. Hieraus ergibt sich für die Digital Humanities ein Methodenrepertoire, das von digitalen Texteditionen in der Mediävistik über soziale Netzwerkanalysen in der Kommunikationswissenschaft bis zu kognitiven Modellierungen in der Psychologie und 3D-Rekonstruktionen in der Bauforschung reicht. Die inhaltliche Breite der Digital Humanities drückt dabei auch ein verstärktes Interesse an interdisziplinären Forschungszusammenhängen aus.

Im Gegensatz zu den Natur- und Ingenieurwissenschaften haben die Geisteswissenschaften in der Breite erst vergleichsweise spät informatische Methoden eingesetzt. Dies schmälert nicht die Bedeutung mancher Pionierprojekte. So wurde der digitale Thesaurus Linguae Graecae bereits 1972 begründet und Geoinformationssysteme können sogar auf eine noch längere Nutzung verweisen. Der späte Zugang zu informatischen Methoden betrifft denn auch weniger die Verwendung digitaler Medien in einzelnen Projekten als die Umstellung auf integrierte digitale Abläufe in der Breite der wissenschaftlichen Arbeit.

Der Hauptgrund für die Ungleichzeitigkeit, mit der sich die Informationstechnologien in den verschiedenen Wissenschaften verbreiteten, liegt in der unterschiedlichen Komplexität der zu lösenden Probleme und den eingesetzten Methoden. In den Natur- und Ingenieurwissenschaften sind die Primärdaten oft Messdaten, die leicht in digitale Form gewandelt werden können und sich ohne Informationsverlust als Tabellen darstellen und statistisch verarbeiten lassen. Die Primärdaten in den Geisteswissenschaften sind dagegen meist Texte, etwa Archivalien, die den Geschichtswissenschaften als Quellen dienen. Daneben treten materielle Kulturgüter und deren digitale Repräsentationen in Form von Bildern, 3D-Modellen oder Metadaten, die kontextualisiert und interpretiert werden müssen. Dieser Gegensatz – strukturierte Daten der Naturwissenschaften, unstrukturierte oder semi-strukturierte Daten der Geisteswissenschaften – erklärt einen Teil der spezifischen Anforderungen der Geisteswissenschaften.

Texte spielen in den Digital Humanities nach wie vor eine zentrale Rolle. Auf der anderen Seite steigt aber die Bedeutung der Dokumentation materieller Kulturgüter sowie sozialer Beziehungen (soziale Netzwerkanalyse, Social-Media-Mining). In einer zunächst methodisch zu vollziehenden, dann informatisch zu realisierenden medienübergreifenden Integration dieser Daten mit herkömmlichen Textdaten in Form von durchgängigen wissenschaftlichen Arbeitsabläufen liegt eine entscheidende Herausforderung für die Digital Humanities. Im letzten Jahrzehnt haben sich diesbezüglich informatische Lösungsansätze eröffnet, von denen zwei stellvertretend erwähnt werden sollen, weil sie auch für Tendenzen der Erweiterung des Betätigungsfelds der Digital Humanities stehen: die semantische Integration von verschiedenen strukturierten Datenquellen mit Technologien des Semantic Web (z. B. Linked-Data-Repositories) und die Social-Web-Technologien als Quelle für Verhaltensdaten sowie als Verbreitungsmedium für wissenschaftliche Inhalte.

Themen in diesem Heft

Unseren Aufruf zur Einreichung von Beiträgen haben wir unter anderem gezielt an Projekte versandt, die vom Bundesministerium für Bildung und Forschung (BMBF) im Rahmen der Ausschreibung zur Förderung von Forschungs- und Entwicklungsvorhaben aus dem Bereich der eHumanities aus dem Jahre 2011 gefördert werden. Insgesamt wurden zwanzig Beiträge für das Ihnen nun vorliegende Themenheft vorgeschlagen, von denen wir leider nur acht annehmen konnten. Natürlich spielt die Text- und Sprachverarbeitung dabei nach wie vor eine große Rolle. Immer stärker in den Blick kommen aber – wie oben bereits angerissen – auch Netzwerkstrukturen, raumbezogene Aspekte sowie andere Medientypen.

In den textbezogenen Geisteswissenschaften kann man bisweilen eine Kluft wahrnehmen zwischen Ansätzen des *Close Reading*, die die sorgfältige Interpretation einzelner Textpassagen in den Vordergrund stellen, und Verfahren, die unter dem von Franco Moretti eingeführten Begriff des *Distant Reading* zusammengefasst werden. Dabei werden große Textmengen quantitativ und statistisch analysiert. In ihrem Beitrag *Content Analysis between Quality and Quantity – Fulfilling Blended-Reading Requirements for the Social Sciences with a Scalable Text Mining Infrastructure* widmen sich Matthias Lemke, Andreas Niekler, Gary S. Schaal und Gregor Wiedemann dieser scheinbaren Kluft. Sie stellen dazu das vermittelnde Konzept des Blended Reading vor und präsentieren den Leipzig Corpus Miner als ein Werkzeug, das Blended Reading in den Sozialwissenschaften ermöglicht. Das Papier führt damit in ein zentrales Anwendungsfeld aus dem Bereich der Digital Humanities ein und zeigt interessante Möglichkeiten für das Datenmanagement und die Datenanalyse auf.

Von besonderem Interesse in den textorientierten Geisteswissenschaften ist oft die Aufdeckung von Argumentationsmustern. Der Beitrag von Constantin Houy, Tim Niesen, Jesús Calvillo, Peter Fettke, Peter Loos, Annika Krämer, Klaas Schmidt, Maximilian Herberger, Iris Speiser, Alfred Gass, Luc Schneider und Tim Philippi zu diesem Themenfeld trägt den Titel *Konzeption und Implementierung eines Werkzeuges zur automatisierten Identifikation und Analyse von Argumentationsstrukturen anhand der Entscheidungen des Bundesverfassungsgerichts im Digital-Humanities-Projekt ARGUMENTUM* und widmet sich diesem Thema. Konzept, Architektur und Umsetzung eines entsprechenden Werkzeuges werden präsentiert und Einblicke in mögliche Anwendungen gegeben. Der Beitrag betrachtet damit einen sehr relevanten Aspekt der Digital Humanities im Bereich der Jurisprudenz und juristischer Textdatenbanken und darüber hinaus.

In der praktischen Arbeit der textorientierten Geisteswissenschaften, beispielsweise den Literaturwissenschaften oder der Theologie, spielt auch die sogenannte textkritische Edition eine zentrale Rolle. In ihrem Beitrag *Differenzanalyse komplexer Textvarianten: Diskussion und Werkzeuge* beschreiben André Medek né Gießler, Marcus Pöckelmann, Thomas Bremer, Hans-Joachim Solms, Paul Molitor und Jörg Ritter den editionsphilologischen Textvergleich am Beispiel von zwei Editionen als eine besonders charakteristische Anwendung in den Digital Humanities. Werkzeuge zum Textvergleich wie etwa das UNIX-Werkzeug *diff* erlauben es zwar, die Unterschiede zwischen zwei Texten darzustellen, für die editionsphilologische Frage der Entstehung von Texten und des Zusammenhangs verschiedener Textvarianten müssen jedoch aufgabenspezifische Editionswerkzeuge bereit gestellt werden, wie sie im Projektvorhaben SaDA: Semiautomatische Differenzanalyse von komplexen Textvarianten entwickelt werden.

Ein weiterer Anwendungsfall von Textanalyseverfahren findet sich in den qualitativ arbeitenden Sozialwissenschaften, in denen Daten aus qualitativer sozialwissenschaftlicher Forschung für eine sekundäranalytische Nachnutzung bereitgestellt werden sollen. Der Beitrag *Die Erschließung von Primärmaterial qualitativer Studien für die Sekundäranalyse als Herausforderung für Sozialwissenschaften und Informatik* von Volker Baethge-Kinsky und Segej Zerr beschreibt diese Anforderungen und skizziert, wie ein Werkzeug zur semi-automatischen Metadatengenerierung für diese Aufgabenstellung zum Einsatz kommen kann.

Die Leistungsfähigkeit computerlinguistischer Verfahren hängt stark vom Korpus ab, auf dem die Lernalgorithmen trainiert werden. Andrea Horbach, Stefan Thater, Diana Steffen, Peter M. Fischer, Andreas Witt und Manfred Pinkal untersuchen in *Internet Corpora: A Challenge for Linguistic Processing*, wie sich unterschiedliche Zusammensetzungen des Korpus auf die Qualität der Wortartenanalyse (Part-of-Speech-Tagging) von Twitter-Nachrichten auswirken.

Im Beitrag „*im Zentrum eines Netzes [...] geistiger Fäden*“. *Erschließung und Erforschung thematischer Zusammenhänge in heterogenen Briefkorpora* beschreiben Vera Hildenbrandt, Roland S. Kamzelak, Paul Molitor und Jörg Ritter einen semi-automatischen Ansatz zur Unterstützung der kollaborativen Verschlagwortung heterogener Briefkorpora. Die manuelle Verschlagwortung wird dabei basierend auf klassischen IR-Verfahren durch die automatische Extraktion vermeintlich sinntragender Stichwörter unterstützt. Im Anschluss erfolgt eine semiautomatische Zuordnung von Schlagwörter zu diesen Stichwörtern. Die Arbeit ist im BMBF-Projekt „Vernetzte Korrespondenzen. Erforschung und Visualisierung sozialer, räumlicher, zeitlicher und thematischer Netze in Briefkorpora“ entstanden und zeigt auf, wie interaktive Werkzeuge die Erschließung spezieller Dokumentarten unterstützen können.

Die beiden weiteren Beiträge zum Schwerpunktthema haben zwar ebenfalls Bezüge zu Texten, sie zeigen aber beispielhaft medienübergreifende Aspekte auf. So treten Probleme der Geometrierverarbeitung in den Digital Humanities natürlich im Zusammenhang mit Fragestellungen der Archäologie oder Bauforschung auf. Gerfrid Müller, Frank Weichert, Denis Fisseler, Michele Cammarosano, Daniel Bachmann und Jörg Günnewig zeigen in ihrem Beitrag über die *Modellierung einer raumbezogenen Datenbank zur Repräsentation und Analyse syntaktischer und semantischer Merkmale von Keilschrifttafeln* aber einen interessanten anderen Anwendungsbereich. Beschrieben wird ein Verfahren, welches Keilschrift auf 3D-Geometriemodellen von Tontafelfragmenten analysiert und so die Tafelfragmente identifiziert, die von der gleichen Tontafel stammen.

Schließlich führen auch musikwissenschaftliche Editionsprinzipien zu neuartigen Anforderungen an das Datenmanagement, welches den Notentext mit seinen Überar-

beitungsstufen, musikwissenschaftliche Anmerkungen und Tondokumente aufeinander beziehen muss. Im Projektbericht *Das Gesamtkunstwerk Oper aus Datensicht – Aspekte des Umgangs mit einer heterogenen Datenlage im BMBF-Projekt Freischütz Digital* stellen Daniel Röwenstrunk, Thomas Prätzlich, Thomas Betzwieser, Meinard Müller, Gerd Szwillus und Joachim Veit eine entsprechende Systemarchitektur vor, die Tondokumente mit Hilfe von referenzbasierter Audiosegmentierung dem Notentext zuordnet.

Zusammenfassung und Ausblick

Die Digital Humanities als wichtiges und herausforderndes Anwendungsfeld für Datenbanktechnologie und Information Retrieval bringen neben fachlichen Aspekten (wie sie in den Beiträgen zu diesem Heft beschrieben werden) auch zahlreiche weitere Fragen hervor: Welche Konsequenzen ergeben sich für Lehrangebote in der Informatik im Allgemeinen und zu Datenbanktechnologie und Information Retrieval im Besonderen? Sind auf Seiten der Informatik und in der Gesellschaft für Informatik entsprechende Überlegungen und Strukturen zur Bündelung der Initiativen zu Digital Humanities sinnvoll?

Im Hinblick auf die Lehre sind in den vergangenen Jahren zahlreiche Studiengänge und Nebenfachangebote zum Bereich der Digital Humanities entstanden. Einen ersten Überblick gab 2011 die Broschüre „Digitale Geisteswissenschaften“.¹

In fast allen Angeboten kommen in unterschiedlich großem Umfang auch Lehrinhalte zum Datenmanagement, zur Datenbanktechnologie und zum Information Retrieval vor. Für die Einordnung der Angebote wurde dabei das in Abb. 1 dargestellte Modell entwickelt.² Es erlaubt eine Aufschlüsselung der Lehrinhalte nach den Bereichen a bis e gemäß der am unteren Rand der Abbildung angegebenen Skala.

Im Rahmen des BMBF-Projektes DARIAH-DE³ wurde zum Bereich d ein sogenannter Informatik-Kern beschrieben⁴. Dieser versucht ein Minimum an erforderlichen Kompetenzen in Informatik und Angewandter Informatik im Umfang von 32 Credit Points zu beschreiben. Diese werden zu gleichen Teilen auf Abstrakte Grundkenntnisse, Skills/Programmieren, Modelle und Modellieren sowie inhaltliche Anwendungen aufgeteilt. Dabei sind algorithm-

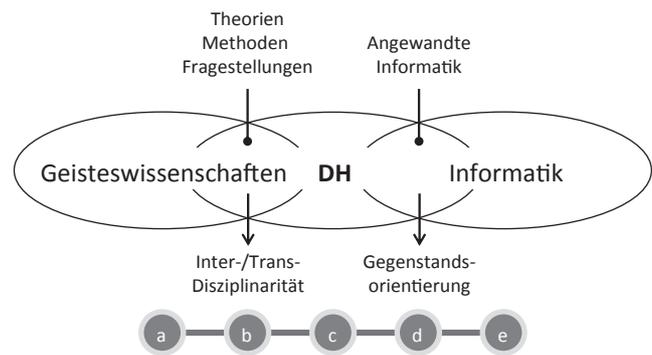


Abb. 1 DH an der Schnittfläche von Geisteswissenschaften und Informatik

mische Grundlagen ebenso vorgesehen, wie Programmierkenntnisse oder X-Technologien. Auch Datenmanagement (einschließlich relationaler und nicht-relationaler Datenmodelle), DBMS und Data-Warehousing werden im Umfang von mindestens 14 Doppelstunden vorgeschlagen.

Studierende in DH-Studiengängen werden daher in Zukunft verstärkt auch Lehrangebote zu Datenbanken und Information Retrieval nachfragen. Ob hierfür spezielle Module sinnvoll sind oder ob eine Nutzung der Angebote für Informatik-Studierende wünschenswert ist, hängt natürlich von vielen Faktoren ab und ist individuell zu klären. Sinnvoll erscheint es aber in jedem Fall, Anwendungen aus dem Feld der DH in der Lehre zu Datenbanken und Information Retrieval zu betrachten. Neben der zunehmenden Bedeutung des Anwendungsfeldes können den Studierenden so auch die verschiedenen Einsatzmöglichkeiten der Datenbanktechnologie an interessanten Beispielen verdeutlicht werden. Wir würden uns freuen, wenn das vorliegende Schwerpunktheft auch hierzu einige Anregungen geben kann.

Im Bereich der Fachorganisationen gründete sich 2013 der Verband *Digital Humanities im deutschsprachigen Raum*⁵ (DHD) als assoziierter Regionalverband der *European Association for Digital Humanities*⁶ (EADH). Auf übergeordneter Ebene ist die EADH wiederum Mitglied in der Dachorganisation *The Alliance of Digital Humanities Organizations*⁷ (ADHO).

Auch auf Seiten der Computerlinguistik und der Informatik haben sich Arbeitskreise zum Thema gebildet. So verfolgt der Arbeitskreis *Computerlinguistik für Digital Humanities* (CL4DH) in der Gesellschaft für Sprachtechnologie & Computerlinguistik das Ziel „die Leistungsfähigkeit computerlinguistischer Werkzeuge und Verfahren für den Einsatz in

¹DARIAH-DE: Broschüre „Digitale Geisteswissenschaften“. Hg. vom CCEH, Universität zu Köln. Erstellt im Rahmen des Projekts DARIAH-DE; gefördert vom BMBF (2011). http://dig-hum.de/sites/dig-hum.de/files/cceh_broschuereweb.pdf

²Sahle, P.: DH studieren! Auf dem Weg zu einem Kern- und Referenzcurriculum der Digital Humanities. DARIAH-DE Working Papers Nr. 1. Göttingen: GOEDOC 2013, URN: urn:nbn:de:gbv:7-dariah-2013-1-5 (2013), S. 28

³<https://de.dariah.eu/>; letzter Abruf 12.12.2014

⁴Nähere Informationen hierzu finden sich auch bei Sahle.

⁵<http://www.dig-hum.de/>; letzter Abruf 4.12.2014

⁶<http://eadh.org/>; letzter Abruf 12.12.2014

⁷<http://adho.org/>; letzter Abruf 12.12.2014

den Digital Humanities zu befördern“⁸. In der Gesellschaft für Informatik hat sich im Dezember 2014 ein Arbeitskreis *Digital Humanities* im Fachbereich *Informatik und Gesellschaft* (IUG) gegründet. Wegen der Bedeutung des Themas wäre es sehr wünschenswert, wenn sich hier auch verstärkt Mitglieder aus dem Fachbereich Datenbanken und Informationssysteme engagieren würden.

Community-Beiträge in diesem Heft

Nachdem Albrecht Blaser im Datenbank-Rundbrief (Ausgabe 16, November 1995) die ersten zehn Jahre unserer BTW-Tagung sehr detailliert beschrieben und nach verschiedenen Kriterien statistisch ausgewertet hatte, haben Theo Härder und Erhard Rahm in einem Community-Beitrag *30 Jahre „Datenbanksysteme für Business, Technologie und Web“ – Die BTW im Wandel der Datenbank-Zeiten* die Aufgabe übernommen, anlässlich ihres 30. Geburtstags ihre gesamte bisherige „Lebenszeit“ zu betrachten. Neben umfangreichen statistischen Auswertungen zum sich wandelnden BTW-Themenspektrum und zu den Beitragsbeteiligungen von Universitäten, Forschungsorganisationen und der Industrie konnte zum ersten Mal auch eine genaue Zitationsanalyse aller BTW-Beiträge seit 1985 durchgeführt werden.

In diesem Heft bietet die Rubrik „Dissertationen“ einen beachtlichen Umfang, da sie zwölf Kurzfassungen von Dissertationen aus der deutschsprachigen DBIS-Community präsentieren kann.

Die Rubrik „Community“ enthält schließlich unter *News* neben einer Kurzdarstellung der beiden neuen Big-Data-Kompetenzzentren in Berlin und Dresden/Leipzig weitere aktuellen Informationen aus der DBIS-Gemeinde.

Künftige Schwerpunktthemen

1 Data Management for Mobility

Mobility is a major factor in our society and daily life. Thus, approaches for data management need to address the resulting dynamics, geospatial and temporal relationships, and distribution of resources. In Web design, the methodology of “mobile first”—developing new Web applications for mobile usage first and adapt it later for the desktop case—is widely embraced by industry. However, it often only considers the user interface and not the data management. This special issue addresses novel approaches and solutions for mobile data management. We invite submissions on original research as well as overview articles covering topics from the following non-exclusive list:

- Data management for mobile applications
- Context awareness in mobile applications
- Analytic techniques in mobile applications
- Management of moving objects
- Data-intensive mobile computing and cloud computing
- Data stream management
- Complex event processing
- Case studies and applications
- Foundations of data-intensive mobile computing

Expected size of the paper: 8–10 pages (double-column)

Issue delivery: DASP-2-2015 (July 2015)

Guest editors:

Bernhard Mitschang, University of Stuttgart

bernhard.mitschang@ipvs.uni-stuttgart.de

Daniela Nicklas, University of Bamberg

daniela.nicklas@uni-bamberg.de

2 Best Workshop Papers of BTW 2015

This special issue of the “Datenbank-Spektrum” is dedicated to the Best Papers of the Workshops running at the BTW 2015 at the University of Hamburg. The selected Workshop contributions should be extended to match the format of regular DASP papers.

Paper format: 8–10 pages, double column

Selection of the Best Papers by the Workshop chairs and the guest editor: April 15th, 2015

Guest editor:

Theo Härder, University of Kaiserslautern,

haerder@cs.uni-kl.de

Deadline for submissions: June 1st, 2015

3 Big Data & IR

The term *Big Data* refers to data and respective processing strategies, which, due to their sheer size, require a data center for the processing, and which become available through the ubiquitous computer and sensor technology in many facets of everyday life. Interesting scientific questions in this regard are the organization and management of Big Data, but also the identification of problems that now can be studied and better understood through the collection and analysis of Big Data. In the context of information retrieval as the purposeful search for relevant content, there are two main challenges:

⁸<http://www.gscl.org/ak-dh.html>; letzter Abruf 12.12.2014

(1) retrieval in Big Data and (2) improved retrieval because of Big Data.

Retrieval in Big Data focuses on the organization, the management, and the quick access to Big Data, but also addresses the creative process of identifying interesting research questions that can only be understood and answered in Big Data. Besides the development of powerful frameworks for the maintenance and analysis of text, multimedia, sensor, and simulation data, an important research direction is the question of what kind of insights Big Data may give us today and in the future.

The second challenge in the context of Big Data & IR is the improvement of retrieval approaches through Big Data. Examples include the classic question of improved Web or eCommerce search via machine learning on user behavior data, the usage of user context for retrieval, or the exploitation of semantic data like Linked Open Data or knowledge graphs.

We are looking for contributions from researchers and practitioners in the above described context. The contributions may be submitted in German or in English and should observe a length of 8–10 pages in the Datenbank-Spektrum format (cf. the author guidelines at www.datenbank-spektrum.de).

Important dates:

- Notice of intent for a contribution: August 15th, 2015
- Deadline for submissions: October 1st, 2015
- Issue delivery: DASP-1-2016 (March 2016)

Guest editors:

Matthias Hagen, Universität Weimar

matthias.hagen@uni-weimar.de

Benno Stein, Universität Weimar

benno.stein@uni-weimar.de