



© Vera Münch/Institut AIFB

„Fachleute, die Semantic Web-Technologien beherrschen, sind momentan noch Mangelware“

Interview mit Professor Dr. Rudi Studer, Institut für Angewandte Informatik und Formale Beschreibungsverfahren (AIFB), Karlsruher Institut für Technologie (KIT) aus Anlass seiner Verabschiedung aus dem aktiven Lehrdienst am 31. März 2017.

Mit seiner über zwanzigjährigen Forschung rund um die Erweiterung und Verbesserung von Websuche mithilfe semantischer Technologien gilt Prof. Dr. Rudi Studer, Institut AIFB am KIT, als einer der Pioniere und Wegbereiter der Semantic Web-Forschung. Er war von 2008 bis 2014 Mitglied im wissenschaftlichen Beirat des FIZ Karlsruhe; hat mit dem FIZ Karlsruhe Projekte zur semi-automatischen Erstellung von Ontologien aus Texten durchgeführt. Von September 2001 bis März 2004 war er außerdem Mitglied im L3S Learning Lab Lower Saxony und dabei für die Abwicklung von Forschungsprojekten sowie die Betreuung von Doktoranden am L3S zuständig. Am FZI Forschungszentrum Informatik hat er viele Jahre als Direktor Forschungserkenntnisse aus der Universität in prototypische Anwendung für Wirtschaft und Gesellschaft transferiert und sich als wissenschaftlicher Vorstand für die Einrichtung eingesetzt. Mit Erreichen der Altersgrenze verabschiedete er sich Ende März 2017 aus dem aktiven Lehrdienst am KIT. Die Journalistin Jennifer Warzecha hat für b.i.t.online ein Gespräch mit ihm geführt.

Herr Professor Studer, Sie gelten als ein „Konstrukteur des Semantic Web“ und stehen damit vorbildlich für den Brückenschlag aus der Wissenschaft in die Wirtschaft und Gesellschaft. Worin genau besteht diese Verbindung?

» **Rudi Studer** ◀ Eine Zielsetzung unserer Arbeit war von Anfang an, eine durchgehende Kette zu bilden, von der Grundlagenforschung über angewandte Forschung, bis hin zum Know-how-Transfer in die Praxis. Das wollten wir für ein breites Themenspektrum tun; für verschiedene Anwendungsfelder, für die Semantische Technologien Bedeutung haben. Zielführend hierfür waren verschiedene Institutionen: einige, die es am Standort Karlsruhe schon gab, andere, die wir neu gegründet haben, um diese durchgehende Kette realisieren zu können. Langjährig bewährte Institutionen sind das FZI Forschungszentrum Informatik am KIT, das eine erfolgreiche Tradition im wissenschaftlich fundierten Know-how-Transfer hat, oder das CyberForum, eines der erfolgreichsten Unternehmer- und Start-up-Netzwerke in Europa. Das Karlsruhe Service Research Institute (KSRI) haben wir am KIT gegründet, um den Know-how-Transfer insbesondere über „Industry-on-Campus“-Modelle zu stärken. All diese Aktivitäten passen auch bestens zur Mission

Semantische Technologien sorgen dafür, dass man als Internet-Nutzer blitzschnell Antworten auf seine Fragen bekommt und weiterführende Informationen ungefragt gleich mit dazu geliefert werden. Firmen wie Google, aber auch kleine und mittlere Betriebe (KMU), profitieren von Studers Absolventen und ehemaligen Doktoranden ebenso wie Universitäten, wo diese als Professoren arbeiten.

Die Mission der Forschenden: Informationen aus unterschiedlichen Quellen wie Bild, Text und Video zusammenzuführen. Eine der Pionierarbeiten der Forschungsgruppen von Studer am Institut für Angewandte Informatik und Formale Beschreibungsverfahren (AIFB) des KIT ist Semantic Media Wiki, eine semantische Erweiterung vom MediaWiki, der Software, auf der Wikipedia läuft. Das sicherlich erfolgreichste in seinen Forschungsgruppen entwickelte Werkzeug ist das „Semantic MediaWiki“. Das Tool wird so gut angenommen, dass es in einer Vielzahl von Anwendungen weltweit im Einsatz ist. Es ergänzt die Software MediaWiki um semantische Erweiterungen. So können Wiki-Seiten mit semantischen Annotationen ergänzt werden. Alle Semantic MediaWiki-Benutzer können damit kollaborativ eine gemeinsame Wissensbasis aufbauen. Auf diese Weise können Abfragen ähnlich zu Datenbanken beantwortet werden (wie z.B. „Welche Semantic Web-Bücher wurden 2016 publiziert?“), oder zusammenfassende Listen (wie z.B. „Die Liste aller gerade in Restauration befindlichen Bücher aus dem 16. Jahrhundert“) automatisch generiert werden.

Semantic Media Wiki ist zugleich die Basistechnologie, die den Wikidata-Ansatz beeinflusst hat. Mit Wikidata* wird gerade ein zentraler Datenpool für Wikipedia aufgebaut, der dann sowohl das chinesische, wie auch das englische und das deutsche Wikipedia beliefern soll. Das Ergebnis wird eine bessere Konsistenz der Daten sein. Der Nutzer erhält beispielsweise, wenn er die Einwohnerzahl von Karlsruhe nachschlägt, stets die gleiche Zahl, egal, welche Landeswikipedia er benutzt.

Semantik MediaWiki wurde von den Wissenschaftlerinnen und Wissenschaftlern um Studer im Forschungsprojekt SemanticCorA (Semantic Collaborative Corpora Analysis for Humanities and Social Sciences) eingesetzt, um eine virtuelle Forschungsumgebung für die kollaborative Analyse von digitalisierten Textkorpora in der Historischen Bildungsforschung zu entwickeln. So können digitalisierte Texte zusammen mit ihren Metadaten quantitativ und qualitativ analysiert werden.

Auch die Verknüpfung von Informationen über Sprachgrenzen hinweg haben die Semantic-Web-Pionierforscher untersucht; z.B. innerhalb des EU-Forschungsprojektes xLiMe (cross-Lingual cross-Media knowledge extraction). Dies führte u.a. zur Erstellung der xLiD-Lexika, die Entitäten wie z.B. Autoren mit ihren umgangssprachlichen Namen in verschiedenen Sprachen verknüpfen, oder zu Ansätzen wie LinkSUM, mit denen aus einer Vielzahl von heterogenen Informationsquellen automatisch die Eigenschaften von Autoren, die besonders relevant sind, für ihre Beschreibungen ausgewählt werden.

* https://www.wikidata.org/wiki/Wikidata:Main_Page

The World's Leading
Library Logistic Partner



Als Partner für automatisierte Bibliothekslogistik beraten wir bei der Planung, der Anlagenkonzeption und der Realisierung

- > UniCar: Schonender Transport
- > UniCar ADAL®: Schnellste Verfügbarkeit der Medien
- > UniSortCar: Transport und Sortierung mit einem System

Telelift GmbH
Frauenstr. 28
82216 Maisach
www.telelift-logistic.com

Im Rahmen des 33. AIK-Symposiums „Von Wissensmanagement zu Web Science“ Ende März in Karlsruhe übergab Professor Dr. Rudi Studer (Mitte) den Staffelstab an seinen Nachfolger Professor Dr. York Sure-Vetter. Im Bild vorne: Ehefrau Irene Studer, ganz links: Professor Dr. Harald Sack, Institut AIFB und FIZ Karlsruhe.



des KIT, die neben Forschung und Lehre eben auch Innovation umfasst.

Das Semantic Web ist ein Kernthema Ihrer Forschung. Was bringt es genau?

» Rudi Studer ◀ Das markanteste Beispiel ist heute, dass man bei der Websuche mit Google oder den Suchmaschinen anderer Anbieter gleich einen Teil der Antwort, nach der man sucht, geliefert bekommt, nicht wie zu Anfangszeiten des Webs einfach nur Links auf Dokumente, die mehr oder weniger relevant waren. Außerdem bestehen die Antworten nicht mehr nur aus Text, sondern es werden auch Fakteninformationen mitgeliefert. Also wenn ich ein Produkt suche, bekomme ich auch noch den Preis oder Lieferfristen mit angezeigt. Ich bekomme die Öffnungszeiten eines Restaurants oder die wichtigsten Angaben zu einer gesuchten Person. Das alles geht nur mit semantischen Technologien im Hintergrund.

Haben Sie auch aktiv zur Verknüpfung Wissenschaft - Wirtschaft beigetragen?

» Rudi Studer ◀ Mein Engagement als Direktor am FZI zum Beispiel hat hier bestimmt einiges bewirkt. Wir waren auch eine der ersten Forschungsgruppen, die weltweit betrachtet, von Anfang an zur Entwicklung von Semantic Web-Methoden und Werkzeugen sowie zur Standardisierung von Beschreibungssprachen beigetragen haben. Standardisierung ist ein ganz wichtiger Punkt, um Methoden und Werkzeuge in die Praxis transferieren zu können. Unsere Gruppe war u.a. daran beteiligt, Webstandards wie die Web Ontology Language (OWL) zu definieren. Mit Standards kann ich ein weltweites Netz von Wissensbasen

aufbauen und deren Inhalte bei Bedarf integrieren. Dies ist ein zentraler Unterschied zu den Expertensystemen der 80er-Jahre, die jeweils Insellösungen darstellten, die nicht miteinander vernetzt werden konnten. Nehmen Sie beispielsweise die sogenannte Linked-Data-Initiative, die weltweit verteilte Informationen verbindet. Da hat man hunderte verschiedene Datenquellen, die alle im gleichen Formalismus daher kommen, in diesem Fall im Resource Description Framework RDF, einem der Webstandards des W3C. Der Trick ist, dass all diese Datenquellen durch entsprechende Links miteinander verknüpft sind. Ich kann eine geographische Information zu Karlsruhe zusammenführen mit einer Wirtschaftsinformation über Karlsruhe und mit einer Politikinformation aus Karlsruhe, obwohl diese Informationen aus drei unterschiedlichen Quellen kommen. Weil sie im gleichen Format vorliegen, kann man sie integrieren. Zu diesem methodischen Ansatz haben wir in Karlsruhe so manches beigetragen. Dass man sich auf die W3C-standardisierten Beschreibungssprachen RDF und OWL geeinigt hat, war einer der Wendepunkte in der Semantic Web-Entwicklung.

Gibt es etwas, was aus Ihrer Sicht den Standort Karlsruhe für den Forschungstransfer besonders attraktiv macht?

» Rudi Studer ◀ Ich glaube, was sich in den letzten Jahren hier deutlich verbessert hat, ist eine Innovationskultur, die Know-how-Transfer und Start-up-Mentalität systematisch fördert. Innovation ist auch Bestandteil der Ausbildung am KIT geworden. Hier gibt es eine Professur für Entrepreneurship, hier werden Innovationsevents organisiert, es gibt entsprechende

Lehrveranstaltungen. Unsere Gruppe am AIFB veranstaltet z.B. in jedem Wintersemester ein Seminar, das sich um dieses Thema kümmert, so dass man bei den Studierenden Bewusstsein schafft, Awareness. Wir haben in Karlsruhe schon viel erreicht, aber es gibt immer Luft nach oben. Von Vorteil ist auch die Start-up-Kultur, die wir am KIT und insgesamt in der Technologieregion Karlsruhe pflegen, unterstützt von Einrichtungen wie dem CyberForum und anderen.

Was hat Sie und Ihre Forschungsgruppen in Ihrer Forschung am meisten vorangetrieben?

Rudi Studer In den letzten 20 Jahren war das eben, die Idee des Semantic Web zu realisieren, die anfangs als eine ziemlich verrückte Idee galt. Unser Ziel war es, schrittweise aufzuzeigen, dass die Idee realisierbar und machbar ist und welchen Mehrwert man mit diesem Ansatz erzielen kann; wir wollten das Web zu einer vernetzten Wissensbasis weiter entwickeln.

Wir haben zahlreiche Veranstaltungen dazu angeboten, um über die Technologien und deren Potential zu informieren. Am FZI waren das zum Beispiel Fördervereinsveranstaltungen. Hier hat man gezielt Ergebnisse der Zusammenarbeit zwischen dem FZI und seinen Partnerunternehmen präsentiert und damit den Dialog zwischen Wirtschaft und Wissenschaft gefördert. Wir haben auch sehr industrieorientierte Tagungen organisiert und versucht, Vertreterinnen und Vertreter aus der Wirtschaft als Teilnehmer zu gewinnen. Ich denke, dass wir es seit gut fünf Jahren wirklich geschafft haben, dass semantische Technologien in der Praxis Fuß fassen und benutzt werden. In jeder großen Firma sind heute Abteilungen vorhanden, die mit diesen Methoden, mit diesen Werkzeugen und mit diesen Formalismen arbeiten, und damit Produkte und Services gestalten.

Paradebeispiel ist schema.org. Das ist ein Vokabular (eine Ontologie) zur semantischen Beschreibung von Informationsinhalten, das Google, Microsoft und andere wie Yahoo gemeinsam entwickeln und pflegen. Es wird unter anderem dazu verwendet, Produktangebotsseiten oder Filmseiten semantisch zu annotieren. So können Suchmaschinen wie Yahoo, Bing oder Google die Produkt- oder Filminformationen passend zur Suchanfrage finden. Schema.org definiert zum Beispiel: Was ist ein Preis? Was ist ein Produkt? Was ist ein Lieferant?. Damit hat man ein Vokabular, mit dem Anbieter entsprechende Produktseiten beschreiben können, damit die Suchmaschinen sie als solche erkennen können. Aufbauend auf dem schema.org-Vokabular können die Produktinformationen in dem Faktenfenster, das als Teil der Suchergebnisse

The screenshot shows the website www.elfi.info with a navigation bar containing the logo and the text 'Willkommen, Herr Dr. Mus' and 'Profil: Eigenes Profil'. Below the navigation bar, there are two main sections: 'Förderprogramme nach I' and 'Förderart (25 / 2951)'. The 'Förderprogramme nach I' section includes a dropdown menu for 'Forschungsgebiete (696)' with options like 'Geistes- und Sozialwis', 'Ingenieurwissenschaft', 'Lebenswissenschaftler', and 'Naturwissenschaften'. The 'Förderart' section includes a list of funding types with checkboxes: 'Beihilfe (60 / 500)', 'Investitionskredite (23)', 'Nachwuchsförderung', 'Preise (495 / 461)', 'Projekt (747)', 'Sonstige (154)', and 'Studie (53)'. To the right of these filters, there is a list of '696 allgemeine u' funding programs, with the first few entries being 'L'Oréal Deutschland - L Roche-Posay', 'VWS - VolkswagenStift', 'Wissenschaftsförderung Deutschen Brauwirtsch', 'Georg Thieme Verlag K neue Programme', 'Stiftung Prof. Joachim I', 'PAJUNK Medizintechn', and 'Dräger Medical AG & C'.

Preise, Stipendium, Doktorandenförderung, Postdoktorandenpendium, Druckkosten, Tagung, Anschubfinanzierung, Verbundprojekt, Tagungsorganisation, Reisekosten, Existenzstudie, Essay-Wettbewerb, Dissertationssubvention, Auslandsförderung, bilaterale

In unserer Datenbank finden Sie die Finanzierung für Ihr Forschungsprojekt!

ELFI Gesellschaft für
 Forschungsdienstleistungen mbH
 Postfach 25 02 07
 D-44740 Bochum
 Tel. +49 (0)234 / 32-22940



© Vera Münch/Institut AIFB

Prof. Dr. Rudi Studer lehrte und forschte 27 Jahre am Institut für Angewandte Informatik und Formale Beschreibungsverfahren (AIFB) des Karlsruher Instituts für Technologie (KIT). Am 31. März 2017 ging er in Ruhestand. Prof. Dr. York Sure-Vetter, Professor für Web Science, führt die Forschungsgruppen am AIFB, am FZI und am KSRI mit ihren Forschungsarbeiten zum Web weiter.

angezeigt wird, von der Suchmaschine selbstständig präsentiert werden. Oder wenn man nach einem Film sucht, bekommt man zum Titel und der Inhaltsangabe im Faktenfenster dann beispielsweise noch den Namen des Produzenten, das Erscheinungsjahr oder andere Informationen angezeigt.

Wenn Sie in die Zukunft blicken: Wann glauben Sie, wird es soweit sein, bis Semantic Web und das Maschinelle Lernen allgemein bekannt sind und die Wirtschaft weiß, welche Chancen sie bieten?

» Rudi Studer ◀ In den großen Firmen ist die Technologie weitgehend bekannt. Sie wird zumindest in manchen Bereichen bereits genutzt. Im Mittelstand hat man da sicherlich noch eine längere Strecke vor sich, den Mehrwert dieser Technologien zu vermitteln, also erst einmal ein Verständnis dafür zu generieren. Im zweiten Schritt kann man das dann in passende Produkte oder Services umsetzen. Da gehen sicher noch mal fünf, zehn oder sogar 15 Jahre ins Land.

Was ist es denn, das die KMUs von der Einführung dieser neuen Techniken abhält: die Angst vor Neuem?

» Rudi Studer ◀ Nein. Aber man benötigt Know-how bei seinen Mitarbeitern, die solche Technologien gelernt haben oder verstehen können. In vielen Firmen gibt es oftmals nicht viele IT-Spezialisten und die Uni-

versitäten können auch nicht beliebig viele Absolventen ausbilden, die mit diesen Technologien vertraut sind. Da herrscht ein Engpass. Ich kann heutzutage immer noch Informatik studieren und habe noch nie etwas von Semantic Web gehört. Das ist ja kein Basis-Kurs wie „Ich lerne Programmiersprachen“. Programmiersprachen hat jeder Informatiker gelernt. Den Umgang mit Semantic Web-Technologien beherrscht nicht jeder Informatiker von vornherein.

Von daher hängt die Durchdringung in der Praxis auch von der Verfügbarkeit von Mitarbeitern ab, die diese Technologien gelernt haben. Und da braucht man einen längeren Atem, als es einem manchmal recht ist. Aktuell herrscht da noch ein Fachkräftemangel.

Gibt es an Ihrem Institut gezielte Maßnahmen, um Studierende für Ihr Thema anzuwerben?

» Rudi Studer ◀ Wir möchten die Studierenden durch gute Vorlesungen für unsere Gruppe interessieren. Wir bieten ihnen am Institut interessante Bachelor-/Masterarbeiten an; auch interessante Jobs für wissenschaftliche Hilfskräfte in unseren zahlreichen Drittmittelprojekten.

Einer Ihrer ehemaligen Doktoranden hat es bis in die Forschungsabteilung von Google geschafft. Sie haben mit Ihrer Arbeit die Wirtschaft und Gesellschaft offensichtlich geprägt. Sind Sie stolz darauf?

» Rudi Studer ◀ Was wir hier am KIT sicher erreicht haben, ist, dass Leute, die bei uns in der Gruppe promoviert haben, verantwortungsvolle, einflussreiche Jobs in der Praxis gefunden haben; sicher nicht alle, aber doch eine beträchtliche Zahl. Damit funktioniert der „Know-how-Transfer über Köpfe“, wie wir am FZI Forschungszentrum Informatik zu sagen pflegen. Die Doktoranden, die bei mir promoviert haben, waren immer gefragt.

Manche meiner ehemaligen Doktoranden haben die Hochschullaufbahn eingeschlagen, sind Professoren an Hochschulen oder Universitäten geworden. Sie tragen dadurch zu einer breiteren Ausbildung Studierender in diesen Technologien bei, indem sie Vorlesungen zum Semantic Web oder Semantischen Technologien anbieten. Damit sind sie wichtige Multiplikatoren für die Idee des Semantic Web.