

Feature

Kulturelles Gedächtnis befragen: HPI hilft Filmsequenzen im Web zu finden

27. Oktober 2009

Potsdam/Berlin. Weltweit explodiert in den Medienarchiven die Zahl digitaler Aufzeichnungen: Das „kulturelle Gedächtnis“ speichert immer gewaltigere Mengen von Informationen. Nur ein Teil des Materials ist derzeit übers Internet aufzufinden. Doch wie kann der wachsende Bestand wichtiger Texte, Bilder, Audio- und Videoaufzeichnungen der Zeitgeschichte für intelligente Internetrecherchen inhaltlich noch tiefer erschlossen und besser aufbereitet werden? Was ist zu tun, damit der Schatz dieses kulturellen Erbes gehoben und für kommende Generationen gesichert wird? Wie schafft man es, das enorme Programmvermögen von vielen Interessenten jederzeit und überall nutzen zu lassen? Mit solchen Fragestellungen beschäftigen sich derzeit Forscher des Hasso-Plattner-Instituts (HPI) am Film- und Medienstandort Potsdam-Babelsberg.

„Mediaglobe“ heißt das vom Bundesministerium für Wirtschaft und Technologie geförderte Projekt, bei dem die HPI-Wissenschaftler eng zusammenarbeiten mit den regionalen Partnern Medien Bildungsgesellschaft Babelsberg gGmbH, defa-Spektrum GmbH in Berlin sowie mit Blue Order Technologies AG in Kaiserslautern. Zum Einsatz kommen sollen vor allem neue Technologien für das „Internet der Dienste“ – unterstützt durch das Forschungsförderprogramm THESEUS.

„Es geht letztlich darum, Film-, Fernseh- und Rundfunk-Archiven dabei zu helfen, ihr audiovisuelles Material optimal zu digitalisieren und umfassend zu erschließen. Das ist wichtig, um die wertvollen Inhalte vielen Interessenten zugänglich und auch wirtschaftlich nutzbar zu machen“, erläutert Prof. Christoph Meinel, Leiter des HPI und des Fachgebiets Internet-Technologien und –Systeme. Dazu seien ganz unterschiedliche Datenbanken und Informationsquellen zu einer Art Informationslandschaft miteinander zu verbinden. Vorhandene Bewegtbilder müssten automatisch analysiert sowie mit erläuternden digitalen Vermerken ergänzt werden können. Fürs exakte Auffinden bestimmter Einzelszenen wird nach Meinels Worten eine intelligente Suchmaschine sorgen.

Entstehen soll also ein universelles, leistungsfähiges Internetportal, mit dem sich in einem großen, verteilten Fundus audiovisueller Materialien bestimmte Inhalte schnell und ohne großen Aufwand finden lassen. Das betrifft sowohl aktuelle wie auch historische Themen. Bei diesem Geschäftsmodell haben die Mediaglobe-Projektpartner einerseits kommerziell interessierte Kunden wie

Lizenzinhaber, Produktionsfirmen, Verlage, Medienarchive und Journalisten im Blick, andererseits Privatleute, die Zeitzeugen sind oder besonderes Interesse an Bildung und Zeitgeschichte haben.

Gesucht: Personen, Orte und Ereignisse in alten Dokumentarfilmen

„Beide Zielgruppen wollen wir mit unterschiedlichen Suchstrategien unterstützen. Der sehr gezielt suchende Profi wird eine spezielle, für ihn schnell erfassbare Benutzerschnittstelle bekommen“, verspricht Wissenschaftler Dr. Harald Sack, der in Meinels Fachgebiet arbeitet. „Und der eher breiter suchenden und forschenden Privatperson helfen wir, indem wir ihr ihrem Suchinteresse entsprechend einen geführten Überblick geben“. Ein erster Anwendungsfall werden die Archive des Mediaglobe-Projektpartners defa-spektrum GmbH in Berlin sein. Sie umfassen unter anderem mehr als 10.000 Stunden dokumentarisches Filmmaterial der letzten Jahrzehnte deutsch-deutscher Geschichte seit den 60er Jahren bis heute.

Für Radio- und Fernseh-Sender sieht das Mediaglobe-Projekt ein weiteres Geschäftsmodell vor. „Deren Redaktionen müssen mit sehr vielen verschiedenen Informationsquellen umgehen, vom Zeitungsarchiv über Sendungsmitschnitte bis hin zu Nachrichtenagentur-Meldungen und Biographien. Oft werden spezifische, höchst unterschiedliche Softwaresysteme eingesetzt, die nur ein kleiner Kreis hoch spezialisierter Fachleute anwenden kann“, beschreibt Sack die Lage. Solche Insellösungen seien aber auch typisch für viele große Unternehmen und Institutionen. Meinel, Sack und ihre Forscherkollegen wollen die verschiedenen Suchtechnologien in einem einzigen Internetportal mit einem einheitlichen Zugang zu allen Quellen zusammenführen. Vor allem vom schnelleren Auffinden und effizienteren Wiederverwenden bereits gesendeter Inhalte versprechen sich die Rundfunk- und Fernsehanstalten hohen Nutzen. Anwendungspartner in dem Mediaglobe-Projekt sind ZDF und Bayerischer Rundfunk – beide Pioniere in der Digitalisierung der Rundfunk- und Fernseh-Produktion.

Automatische und benutzerunterstützte Suchverfahren

Voraussetzung für die erfolgreiche Anwendung solcher Informationssysteme sind zunächst Fortschritte bei der tieferen Erschließung und wirksamen inhaltlichen Analyse des audiovisuellen Materials. Dazu braucht es automatische, vom Benutzer unterstützte Verfahren der semantischen Analyse. „Im Mittelpunkt stehen Techniken, die bei der Erkennung und Identifikation von Personen und Objekten helfen sowie Ereignisse bestimmen und Inhalte zeitlich und geografisch zuordnen können“, berichtet HPI-Forscher Sack. Die Software muss zudem auch natürliche Sprache und Schrift analysieren können.

Was die Strukturierung und Erschließung der audiovisuellen Datenbestände anbetrifft, wird der Benutzer später einmal übers Internet mitwirken. Er kann

dann an entsprechenden Stellen ergänzende Informationen und Kommentare hinzufügen, aber auch Querverweise zu anderen Dateien anlegen, welche die Suche erleichtern. Spezielle Formen der Wissensrepräsentation, mit denen sich die Bedeutung der Informationen formal und maschinenverstehbar ausdrücken lassen (Ontologien), sollen die automatische Verknüpfung mit externen Datenbankinformationen und eine echte semantische Suche ermöglichen. Diese Technologie des „Semantic Web“ hilft dabei, die Eingabe natürlicher Sprache zu verarbeiten, die inhaltliche Bedeutung gestellter Fragen besser zu verstehen und als Antworten ganze Aussagen zu liefern. Im Gegensatz dazu zeigen herkömmliche Suchmaschinen nur Listen von Dokumenten an, welche die Zeichenfolge gesuchter Stichwörter enthalten.

Eine besondere Neuerung besteht künftig in der Implementierung einer „explorativen Suche“. Sie verspricht dem Benutzer durch individuelle Vorschläge und die Aufdeckung verborgener Zusammenhänge ein neuartiges Suchergebnis. Denn oft weiß der Benutzer gar nicht so genau, was er eigentlich sucht, sondern er möchte lieber etwas herumstöbern und hofft, dabei auf etwas Interessantes zu stoßen.

Im Herbst 2011 soll das Mediaglobe-Projekt abgeschlossen sein. „Wir setzen auf den im THESEUS-Programm entwickelten Kerntechnologien auf, und nutzen z.B. verschiedene Verfahren der Video- und Audio-Analyse sowie der automatischen Erzeugung von Metadaten“, erläutert Sack. Das Mediaglobe-Projekt hat einen Umfang von über zwei Millionen Euro.

Kurzprofil Hasso-Plattner-Institut

Das Hasso-Plattner-Institut für Softwaresystemtechnik GmbH (HPI) in Potsdam ist Deutschlands universitäres Exzellenz-Zentrum für IT-Systems Engineering. Als einziges Universitäts-Institut in Deutschland bietet es den Bachelor- und Master-Studiengang „IT-Systems Engineering“ an – eine praxisnahe und ingenieurwissenschaftlich orientierte Alternative zum herkömmlichen Informatik-Studium, die von derzeit 450 Studenten genutzt wird. Insgesamt gut ein Dutzend Professoren und über 50 weitere Lehrende sind am HPI tätig. Es betreibt exzellente universitäre Forschung – auch für erste Adressen der Wirtschaft. Vor allem geht es um Grundlagen und Anwendungen für große, hoch komplexe und vernetzte IT-Systeme. Das HPI kam beim jüngsten CHE-Hochschulranking unter die besten vier Informatikstudiengänge im deutschsprachigen Raum, die sich Rang 1 teilen.

Pressekontakt: Hans-Joachim Allgaier, Tel.: 0331 55 09-119, 06081 57 76 30,
Mobil: 0179 267 54 66, Fax: 06081 96 25 17Mail: presse@hpi.uni-potsdam.de,
info@allgaiercommunication.de

Fotos, Illustrationen und Logos: Katrin Augustin, HPI, Tel.: 0331 55 09-150,
Fax: 0331 55 09-169, Mail: katrin.augustin@hpi.uni-potsdam.de