

# **METADATEN AUS DER CLOUD**

Technologien und Anwendungsfälle der  
Medienschließung mittels Software as a Service

# METADATEN AUS DER CLOUD

## Technologien und Anwendungsfälle der Medienerschließung mittels Software as a Service

**M. A. Michael Eble**  
**Dipl. Inf. Sebastian Kirch**

Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS  
Schloss Birlinghoven | 53754 Sankt Augustin | Deutschland

Einreichung zur 6. Konferenz der Zentralbibliothek Forschungszentrum Jülich (WissKom 2012) in Jülich

Version 1.0      28.02.2012

## **Metadaten aus der Cloud: Technologien und Anwendungsfälle der Medienerschließung mittels Software as a Service**

Das Erschließen von Medienbeständen ist eine zentrale Grundlage, um ihre Inhalte und darin eingeschriebenes Wissen nutzen zu können. Das Erzeugen von reichhaltigen Metadaten gehört daher zu einer wesentlichen Tätigkeit von Bibliotheken und Archiven, um Inhalte in Beständen zu kontextualisieren. Dazu werden manuelle und automatische Verfahren kombiniert und deren Ergebnisse miteinander vernetzt. Insbesondere zur effizienten automatischen Analyse von großen Volumina digitaler Audio-/Video- sowie Bild-/Dokumentenbestände sind leistungsfähige Hard- und Softwarearchitekturen notwendig. Diese sind meist mit hohen fixen Kosten verbunden.

Ansätze des Cloud Computings propagieren demgegenüber nun, Dienste zur Medienerschließung so bereitzustellen, dass eine Verwendung bei weitgehend variablen Kosten möglich wird. Bibliotheken und Medienarchive erhalten damit die Option, leistungsfähige Analyseverfahren individuell und „on demand“ zu nutzen. Über Dienstplattformen wird eine solche „Software as a Service“ bedarfsgerecht verfügbar – ohne, dass zur Projektdurchführung eigene Hardware- und Software-Infrastrukturen lokal vorgehalten werden müssen. Über Schnittstellen sollen Verfahren und Ergebnisdaten in eigene Anwendungen und Archive integriert werden (können). Dadurch können die finanziellen, organisatorischen und technischen Hürden sinken, um automatische Verfahren der Medienerschließung zu nutzen.

Ein Anwendungsfall dessen ist die Verwendung der CONTENTUS-Dienstplattform durch die Deutsche Digitale Bibliothek und ihre Kultureinrichtungen. Die Dienstplattform wurde im Rahmen des vom BMWi geförderten THESEUS-Forschungsprogramms entwickelt und bündelt eine Reihe von Verfahren der Medienerschließung. Kultureinrichtungen und Digitalisierungsdienstleister nutzen die Plattform des Forschungsprojekts, um Digitalisate aufzubereiten sowie strukturell und inhaltlich zu erschließen. Verfahren wie das Entfernen von Alterungseffekten, eine Texterkennung (OCR) oder das Erkennen von in Texten genannten Personen stehen nicht nur via Webbrowser, sondern auch über Web-Service-Schnittstellen zur Verfügung. Bibliotheken und andere Kultureinrichtungen übernehmen auf diese Weise die erzeugten Metadaten aus der Cloud in ihre eigenen Anwendungen.

Der Vortrag adressiert Technologien und Anwendungsfälle des Erschließens und Vernetzens von Wissen aus großvolumigen Medienbeständen. Er basiert auf Marktuntersuchungen und Projektarbeiten zur Medienerschließung via Cloud Computing.

## Referenten



### Michael Eble

Michael Eble, M. A., ist seit 2008 am Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS in Sankt Augustin tätig. Mit einem Schwerpunkt in den Medien- und Kulturbranchen ist er zuständig für Business Development und Produktentwicklung.

Als Mitarbeiter in den Forschungsprojekten CONTENTUS (THESEUS-Forschungsprogramm) und MediaGrid beschäftigt er sich mit Technologien und Anwendungsfällen der Medienschließung mittels Software as a Service.

Telefon +49 22 41 / 14 34 06  
Fax +49 22 41 / 14 4 34 06  
E-Mail [Michael.Eble@iais.fraunhofer.de](mailto:Michael.Eble@iais.fraunhofer.de)



### Sebastian Kirch

Sebastian Kirch ist am Fraunhofer IAIS als Leiter des vom BMBF geförderten Projektes „MediaGrid“

([www.mediagrid-community.de](http://www.mediagrid-community.de)) tätig. Ziel des Projektes ist der Aufbau einer leistungsfähigen Infrastruktur zur Analyse und Verteilung umfangreicher Medienbestände.

Er befasst sich vor allem mit der Konzeption und Implementierung von verteilten Anwendungen im Grid- und Cloud-Bereich, insbesondere der Workflow-Steuerung in Serviceorientierten Architekturen.

Telefon +49 22 41 / 14 19 71  
Fax +49 22 41 / 14 4 19 71  
E-Mail [Sebastian.Kirch@iais.fraunhofer.de](mailto:Sebastian.Kirch@iais.fraunhofer.de)