

Prozessautomatisierung durch Generische Content Transformation. (GCT)

Michael A. Herzog

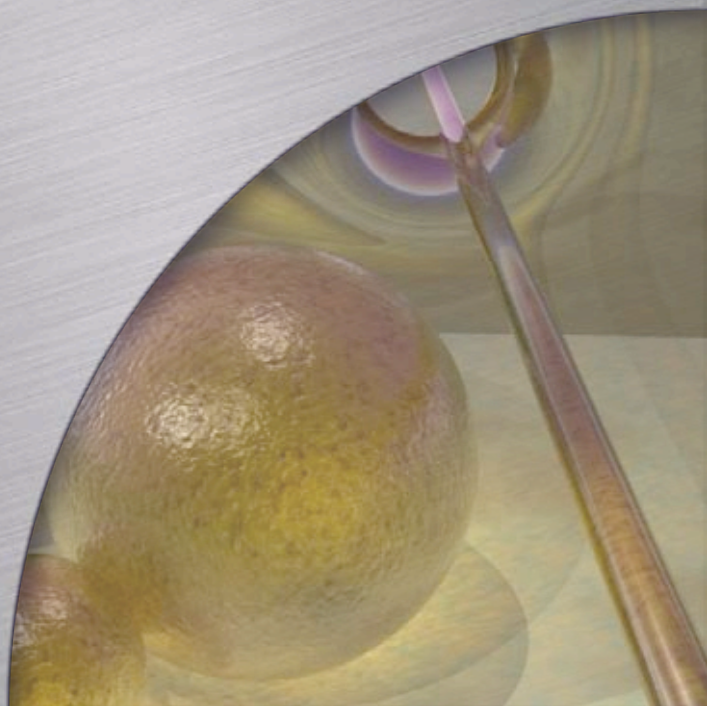
FHTW Berlin/TU Berlin

<http://inka.fhtw-berlin.de/herzog>

IRT Symposium IPTV • 13. März 2007

BACKGROUND

1



GCT-KONZEPT



2

CASE M-LEARNING

3



MEDIA REPOSITORY



4

BACKGROUND

1

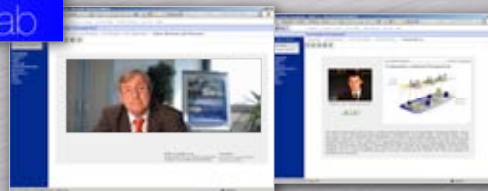


BACKGROUND

- Forschung und Lehre in der Wirtschafts- und Medieninformatik, speziell Media Content Produktion und -Management
- Projektgruppe »vbc.studiolab«
Technische Universität + FHTW Berlin
- Content-Entwicklung für MBI-/CS-Kurse



Virtual Global University



VGU Virtual Global University
School of Business Informatics

www.vg-u.de



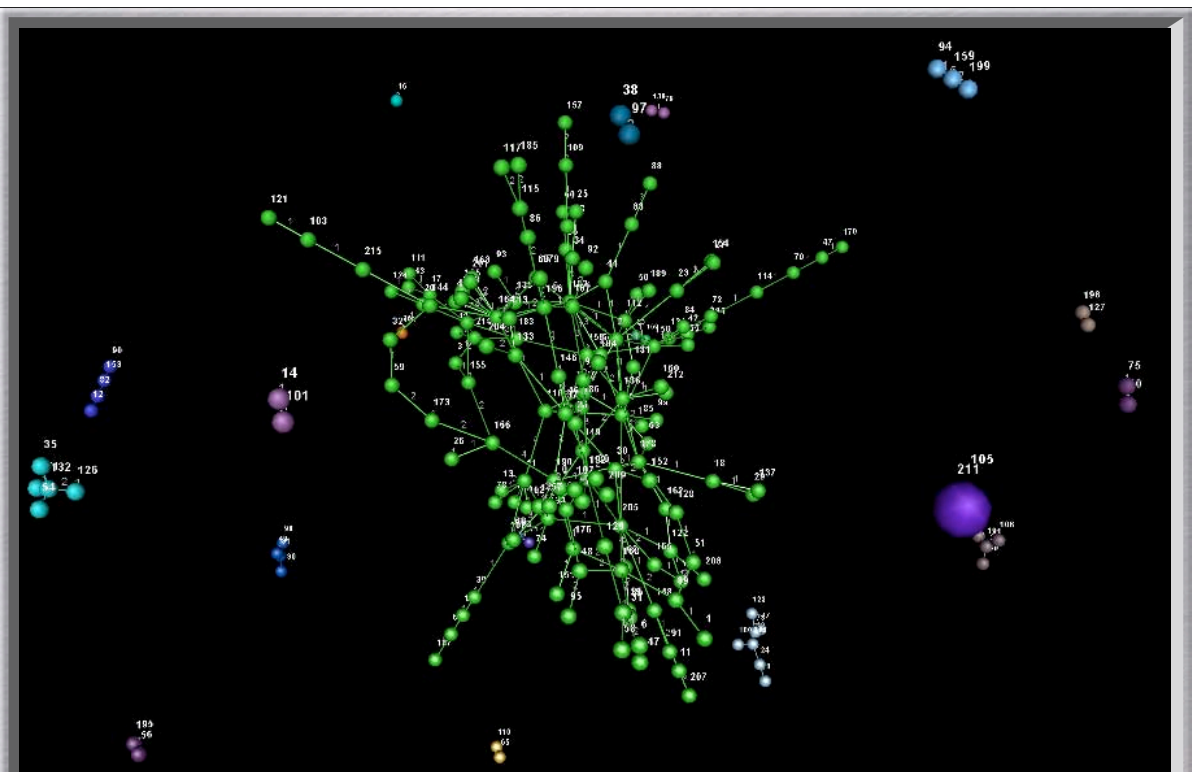
IKAROS

thtw.

Fachhochschule für Technik
und Wirtschaft Berlin
University of Applied Sciences



<http://www.IKAROS-PROJEKT.de>

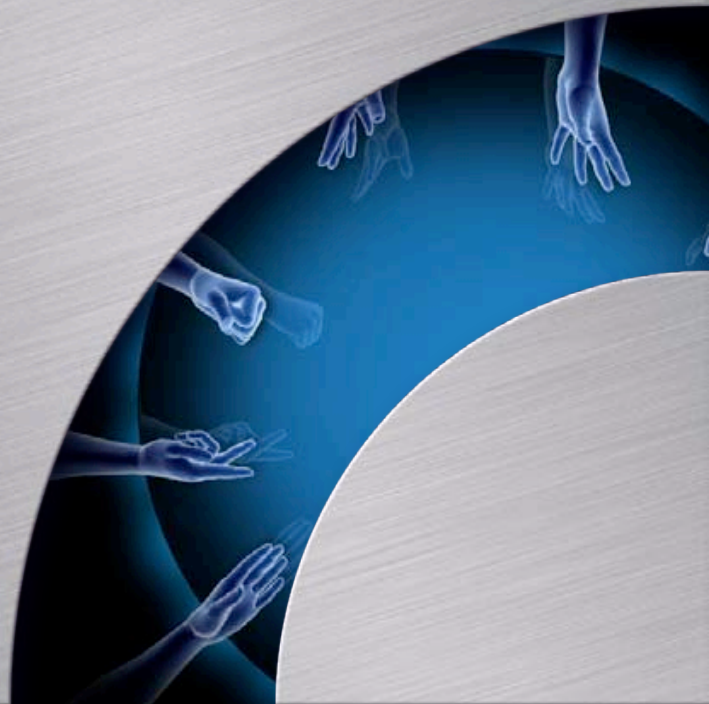


<http://www.commetrix.de>

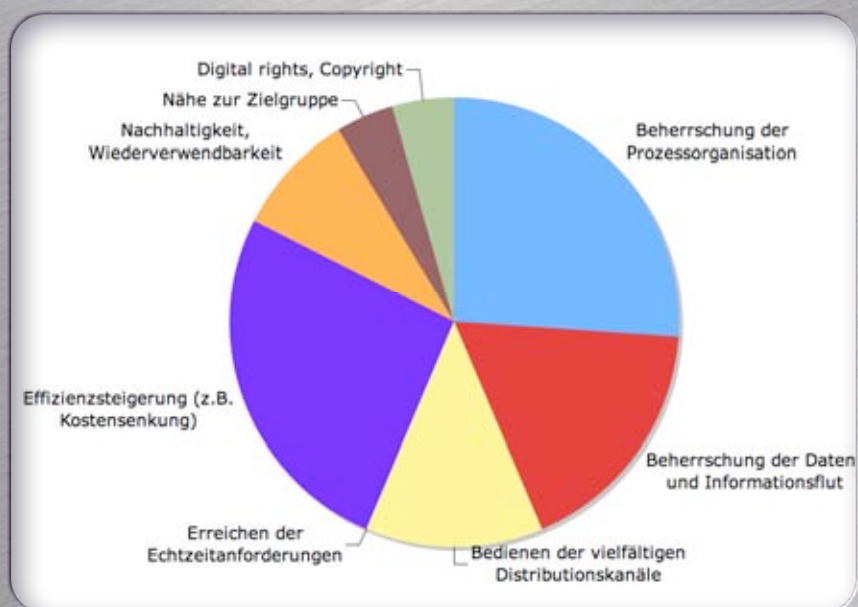


GCT-KONZEPT

2



Herausforderungen in der Medienproduktion



ILM-Studie TUB/FHTW Berlin Stand 2/2007 Interimsergebnisse (to be published)

Content Overload

- Content-Zuwachs im Unternehmen
 - zwischen 60 und 200% pro Jahr¹
- Content ROI
 - Unternehmensdaten: 20% strukturiert / 80% unstrukturiert¹
 - Nur 10% der strukturierten Daten in indizierten Repositories¹
 - Bis zu 40 % Arbeitszeit: Suche nach Inhalten, Behebung von Problemen mit Dokumentversionen, Dokumentbesitz und Umformatierung¹
 - Viele teure Inhalte zu wenig genutzt oder müssen neu erstellt werden.
- Compliance
 - Gesetzliche Verpflichtung zur Datenvorratshaltung
- Collaboration
 - Regionale Datensilos verhindern gemeinsame Nutzung durch verteilt arbeitende Teams. Reduzierte Produktivität. Long Time-to-Market.
- Konsolidierung
 - Eine Infrastruktur für sinkende Kosten, höhere Sicherheit
 - Leistung der Repositories heute: 1 Milliarde Objekte bei $t_r < 45\text{ms}$.

¹) <http://www.jupiterresearch.com>, 4/2006

Trends 2007

- Auffinden digitaler Informationen:
»Google/Spotlight your Intranet«
- Vereinfachung der Erstellung und Verteilung von Content
- Integration unterschiedlicher Applikationen und Daten unter einer personalisierten, prozessorientierten Sicht
- Geschäftsprozess-Integration
- Realtime-Collaboration
- Portalarchitektur:
 - horizontale Services (Directory Service, Enterprise Search, BPM Workflow) und
 - vertikale Services (Authoring, DRT, Collaboration)
- Konzept "Virtuelles Content Repository"²



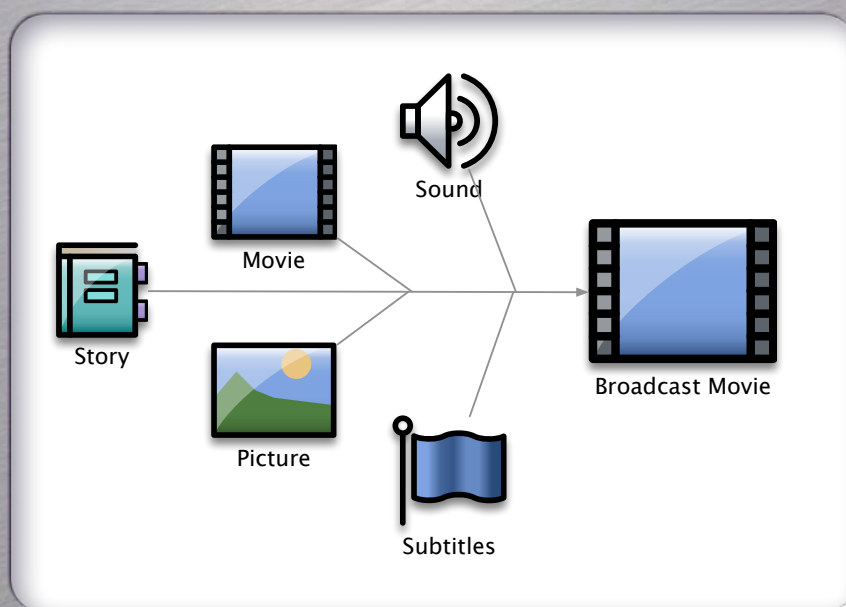
²) Neuwirth, E.: KM im Wandel der Zeit - Vom reinen ECM zum ECCM und Knowledge Management, Oracle Corporation 10/2005, Beitrag auf der Knowtech 2005

Media Content Transformation

- Bedarf
 - ▶ Zahllose Konvertierungsprozesse
 - ▶ Elektronischer Medienbruch!
 - ▶ Viele Werkzeuge

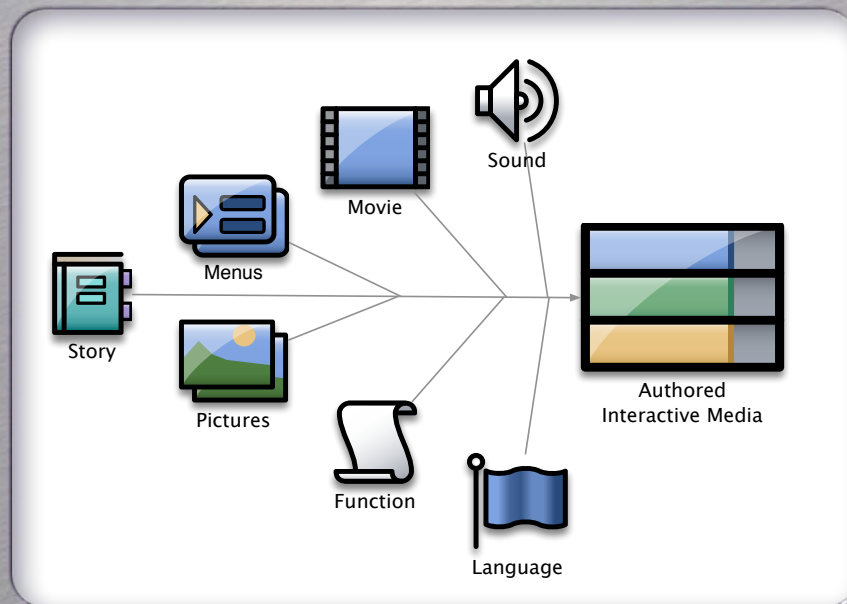


Abgrenzung



Broadcast Movie Production

Abgrenzung



Interactive Media Production

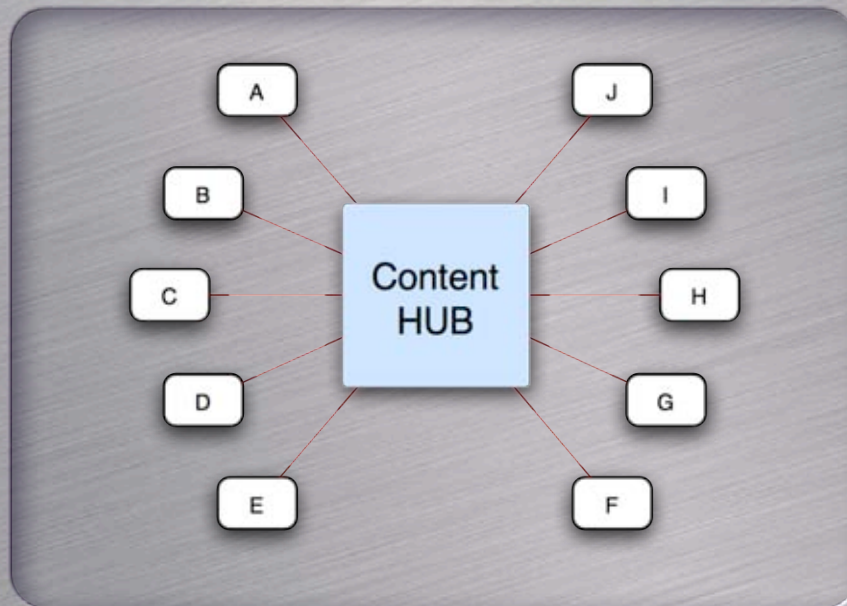
Forschungs- und Entwicklungsschwerpunkt

Konzept der Generischen Content Transformation (GCT)



- ▶ Multimedia Content aus Autorensystemen basiert auf gemeinsamen Prinzipien
- 1. Text und Medienassets sind framebasiert in Vektorrepräsentationen eingebettet
- 2. Interaktive Medien nutzen eine begrenzte Anzahl von gleichartigen Funktionen für die Navigation und Interaktion
- 3. Zeitbasierte Medienassets sind mit den gleichen Steuerfunktionen eingebettet
- 4. Weitere Funktionalitäten sind in strukturierter Textcodierung als Softwareroutinen (z.B. Scripts) abbildbar

Generic Transformation für containerbasierte Medieninhalte (GCT)



GCT-Projekt 1: Rojo-MHP Transformer



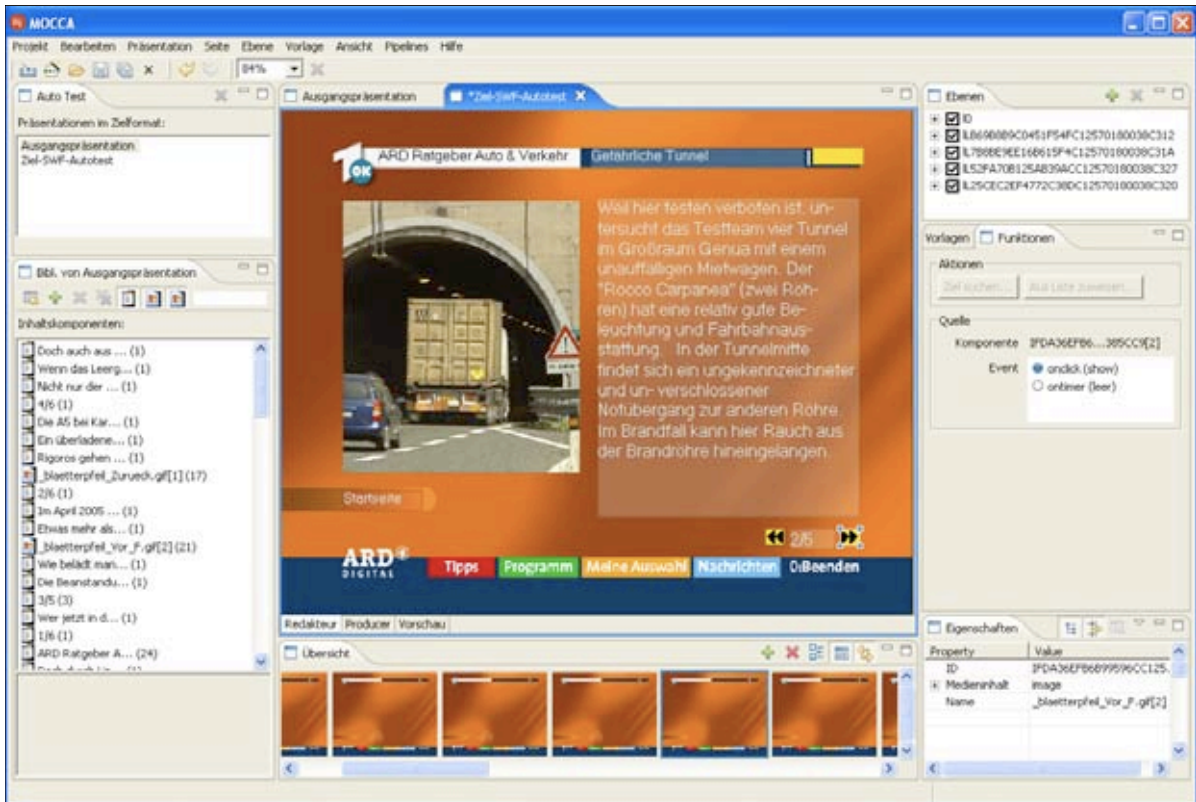
mhp content



RoJo MHP Transforming & Authoring 12/2005

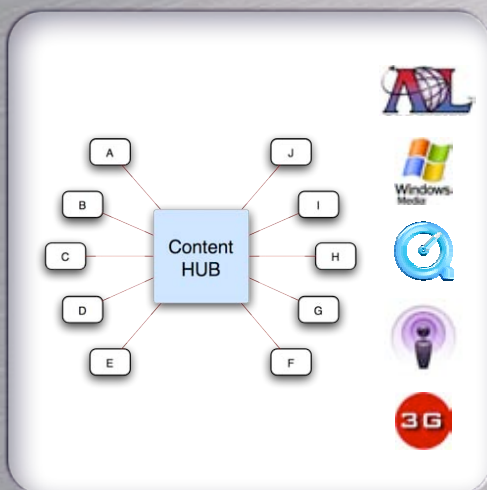


- Content-Bearbeitung und Transformation
- Text und Bild-Layout
- Werkzeug auf XML-Standards



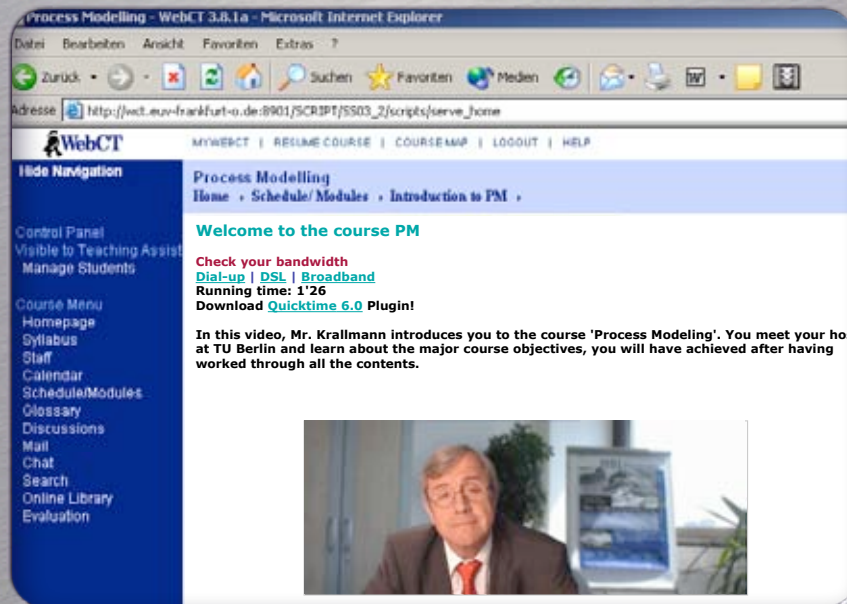
<http://www.moccaonline.de>

GCT-Projekt 2: Rapid E-Learning



- GCT Project rev. 2
 - Rapid-E-Learning-Inhalte-Produktion
 - Import PPT, OOP, KEY, Audio, Video
 - Export HTML, SCO (ADL SCORM)
 - RSS/Podcast
 - Integration der Mediendistribution (FTP, RSS, iTunesU)
 - Ausbau der Medienintegration (Audio, ggf. Video)
 - Simple-Pipeline Implementierung
 - Release 7/2007

Project vbc.studiolab



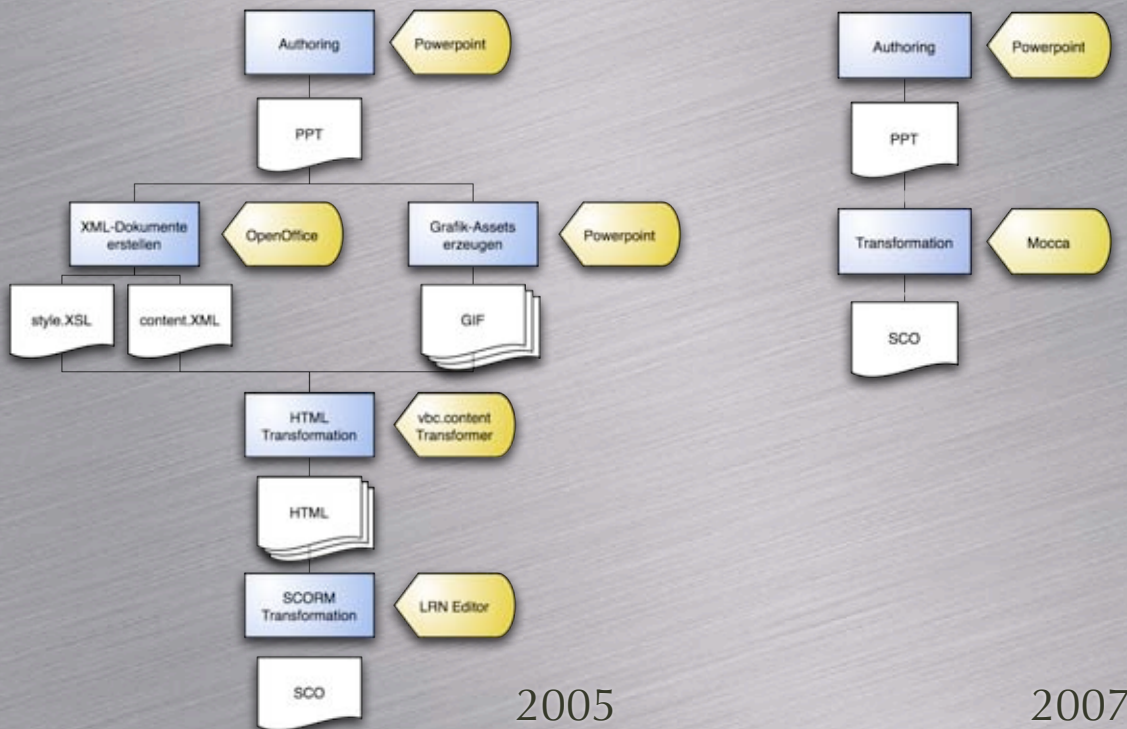
resulting course presentation

ODP online Transformer

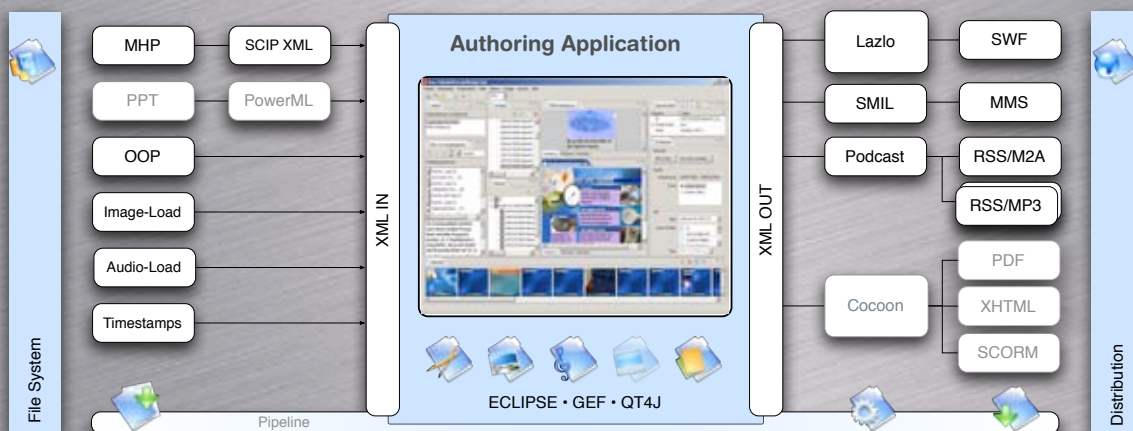


<http://www.moccaonline.de>

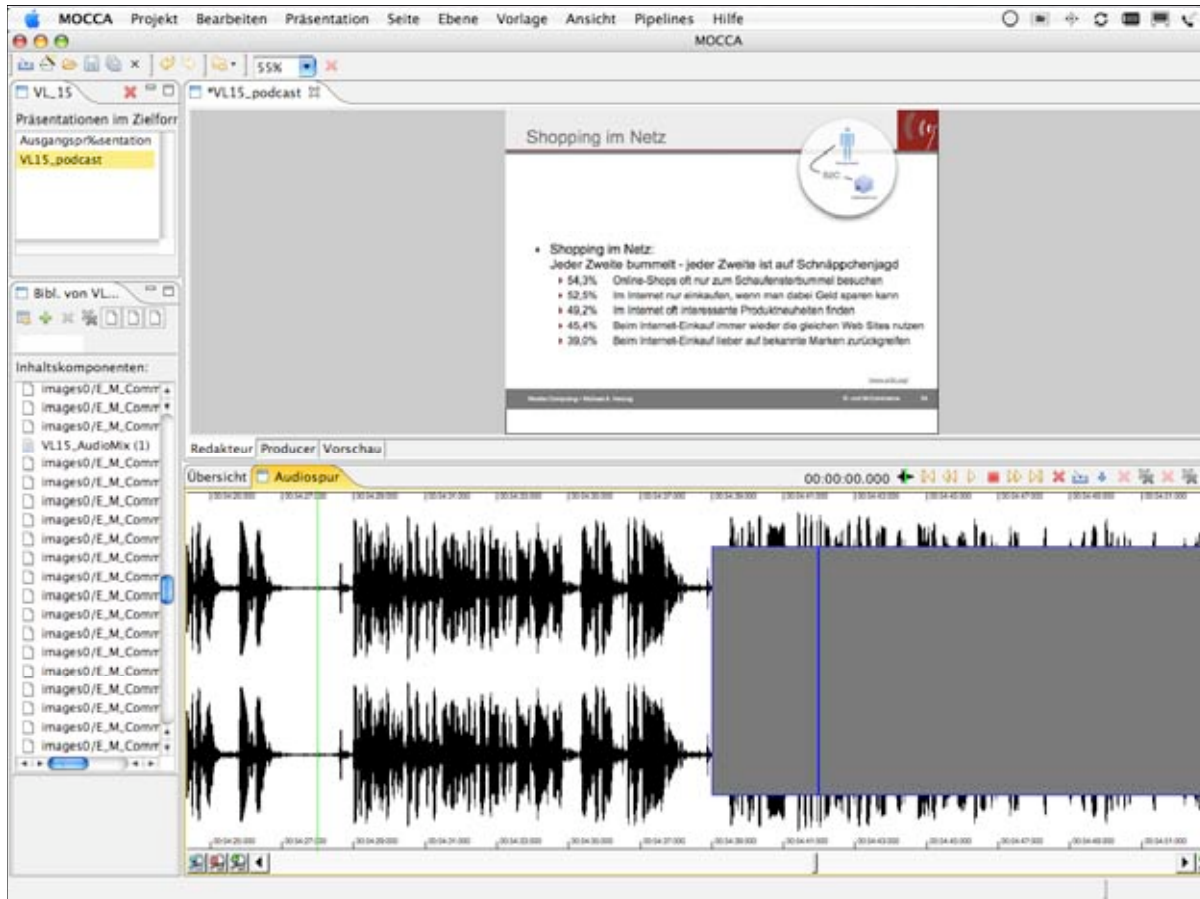
SCORM Produktionsprozess



MOCCA architecture



<http://www.moccaonline.de>

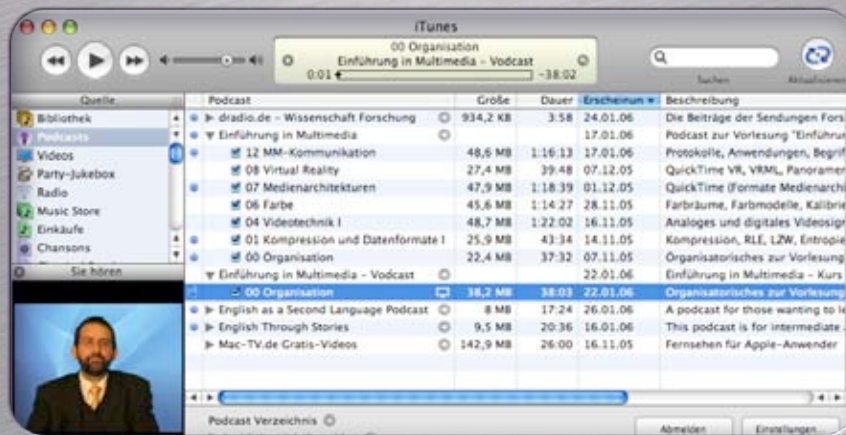


CASE M-LEARNING

3



Mobile Learning - Podcast konsumieren



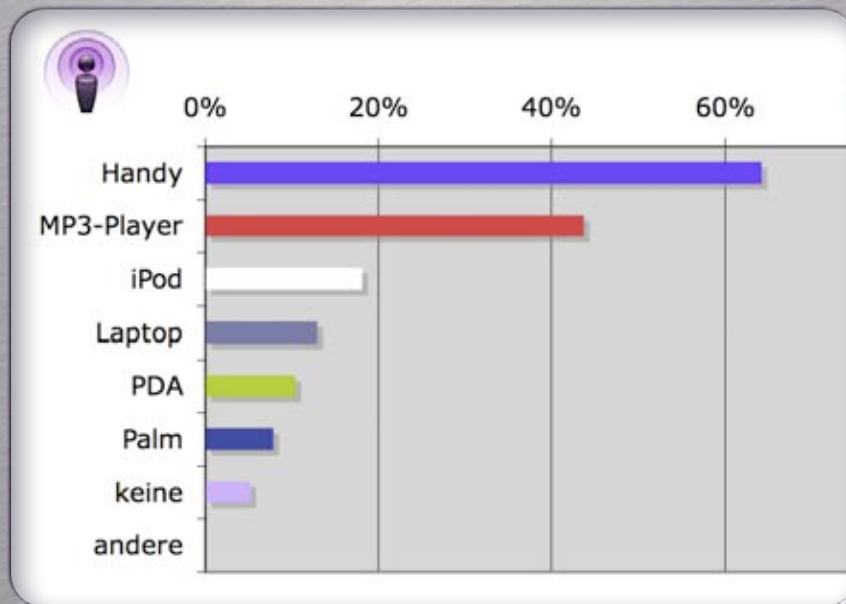
Ansicht eines RSS-Vodcast
im Podcatcher iTunes

Learning on the go



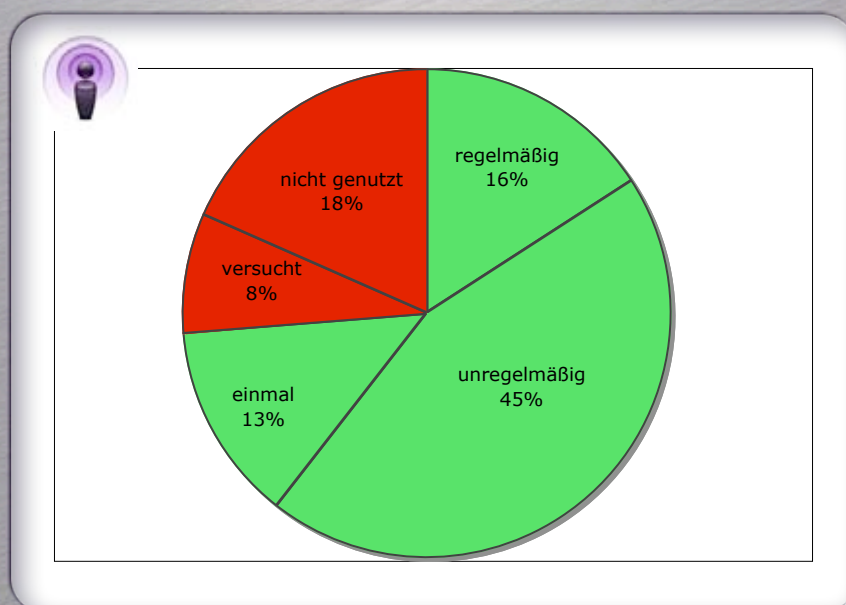
Podcast | Vodcast

Learning on the go



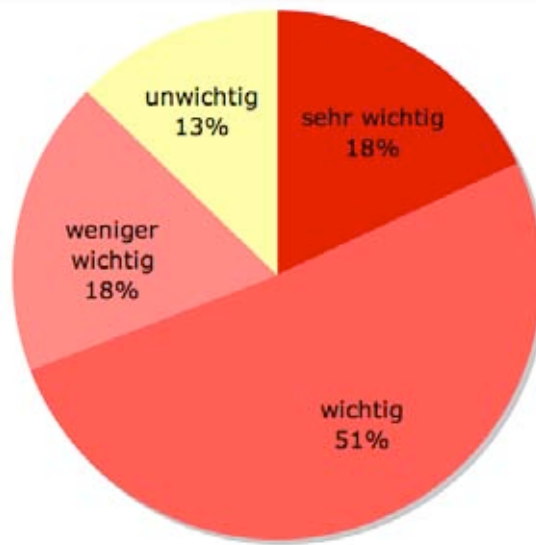
Nutzung mobiler Endgeräte

Learning on the go



70% der Teilnehmer haben das Angebot genutzt.

Learning on the go



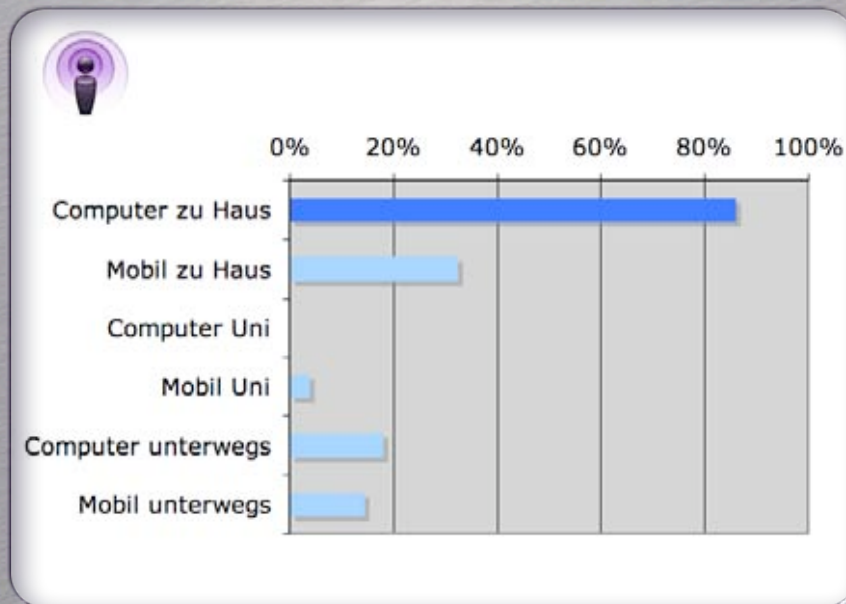
69% der Teilnehmer bewerten das Angebot als wichtig.

Learning on the go

- Eine Podcast-Vorlesung wurde bis zu 3 mal konsumiert
- Podcast-Hörer nutzten durchschnittlich 6 h des Gesamtangebots (16 h Material gesamt)
- Vorlesungen wurden durchschnittlich 1 h am Stück (ohne nennenswerte Unterbrechungen) gehört

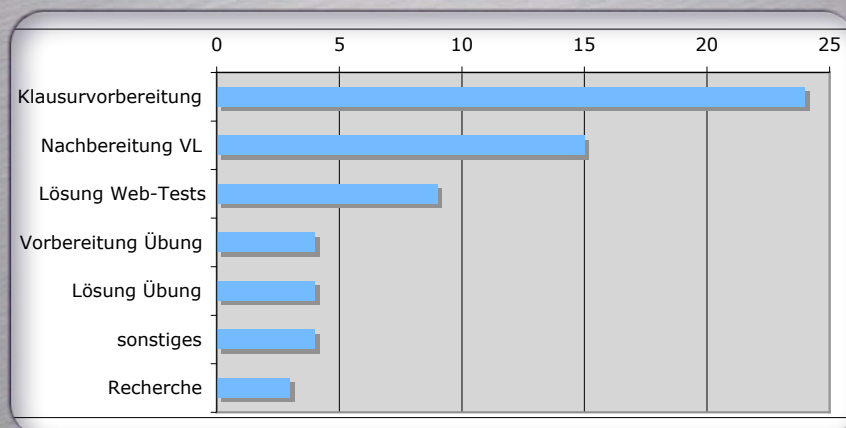


Learning on the go



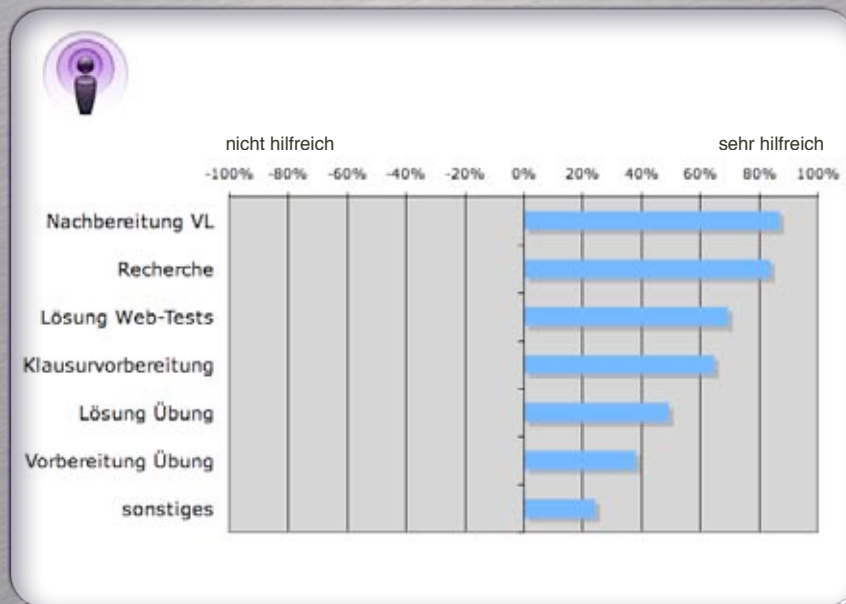
Ort der Rezeption

Learning on the go



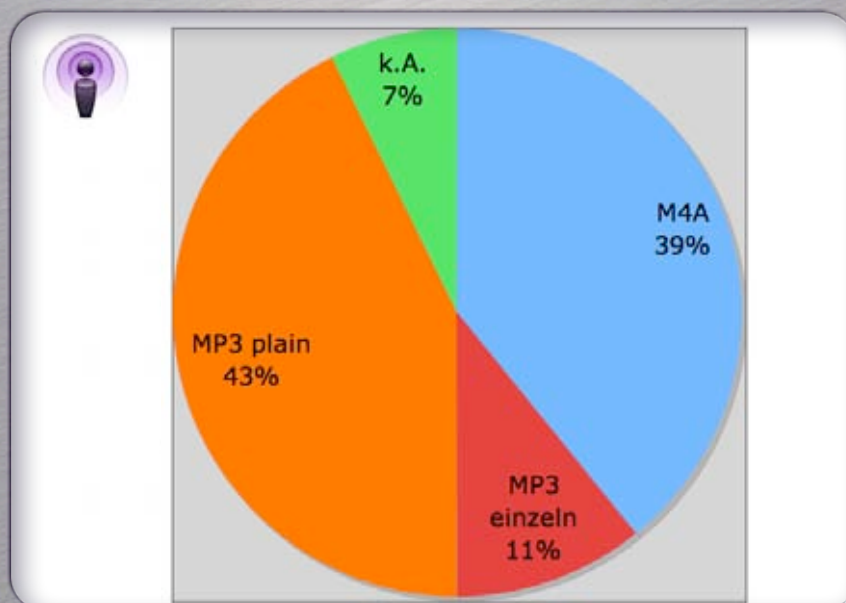
Zweck der Nutzung

Learning on the go



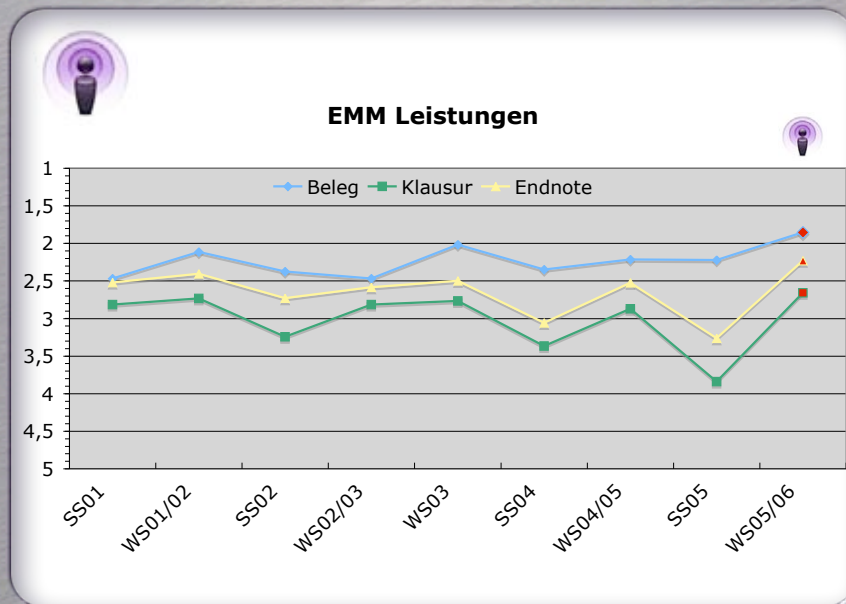
Nutzwert

Learning on the go

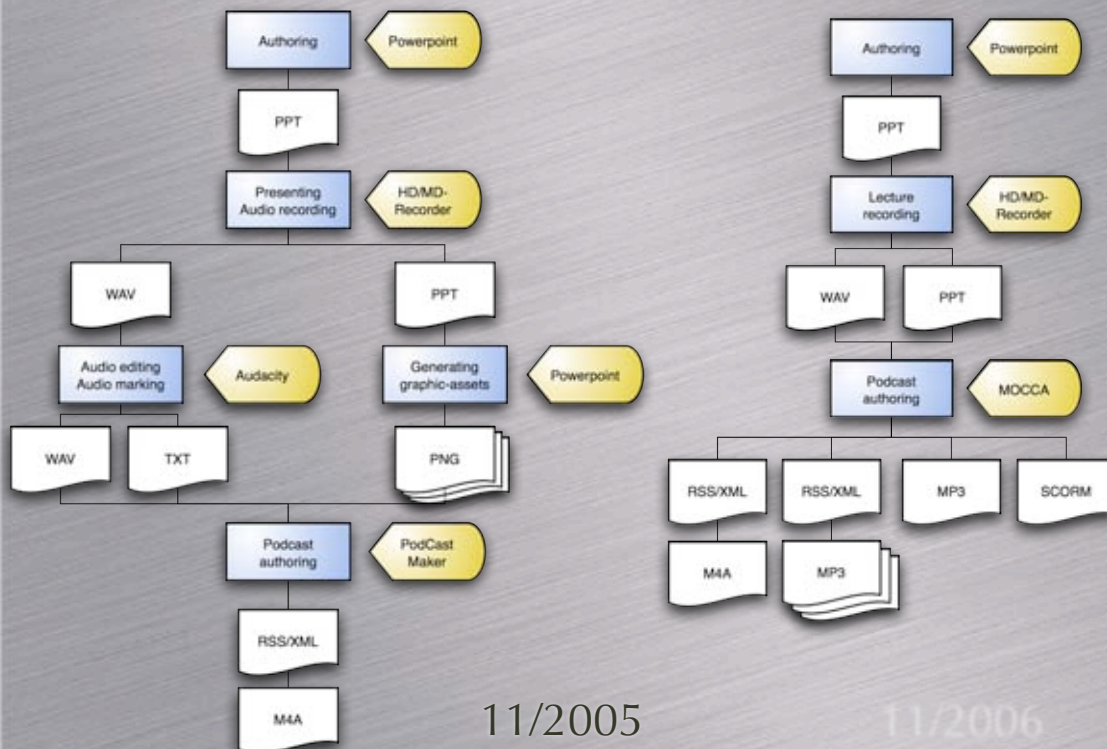


Podcast | Vodcast

Learning on the go



Podcast Produktionsprozess

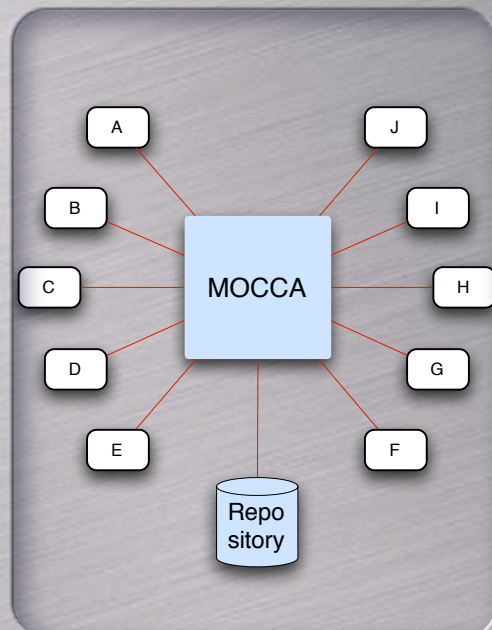


MOCCA MEDIA REPOSITORY

4



GCT-Projekt 3: Content Repository

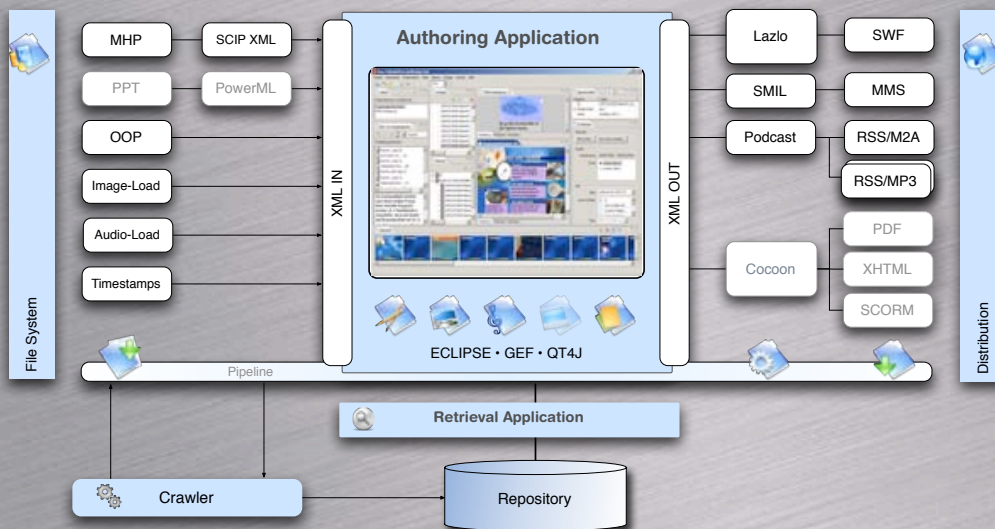


Generisches Datenformat M-XML



- Abbildung vektorbasierter Medienformate aus verschiedensten Projektdatenformaten
- Anlehnung an SVG 1.2
- Unterstützung zeitbasierter Medien
- 80:20 Paradigma

Media Content Repository

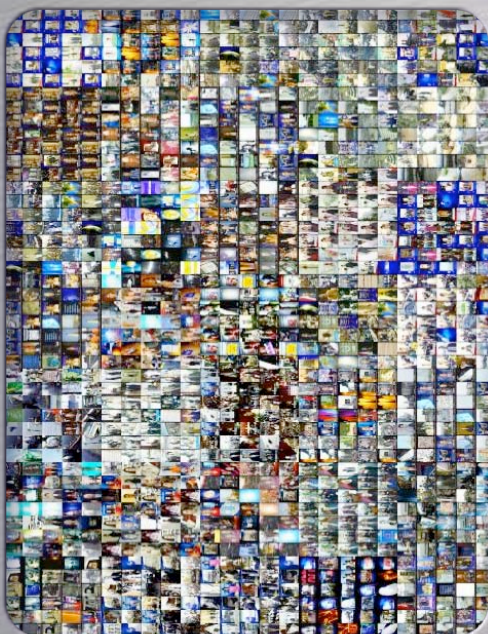


MMR • Retrieval-Konzepte

- Suche nach Assets und Projektdaten
 - Nutzung der Volltext-Daten aus Authoring-Projekten
 - Basis: Generische XML Datenstruktur
 - Vollautomatische Transcodierung in das Enterprise Repository (Crawler)
- CBR für das Auffinden von
 - Asset-Varianten (andere Auflösung, Bildausschnitt)
 - Ähnliche Bild- und Tondateien

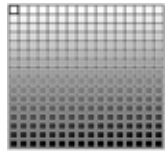


Herausforderung des CBR



- Flut an digital verfügbaren Multimediadaten
 - Digitalfotos, Vektorgrafiken, XML-Dokumente, Filme, Audio-Daten, ...
- Problem: effektive inhaltsbasierte Suche
 - Bsp: Suche Fotos von einem Eisberg
 - Bsp: Suche Fotos vom IRT-Gebäude
 - Bsp: Suche Audioaufnahmen mit dem Sprecher „Dietrich Sauter“
- Datenbankabfrage: SELECT ... FROM ... WHERE ... auf Grund fehlender Attributwerte ungeeignet

Feature-Extraktion

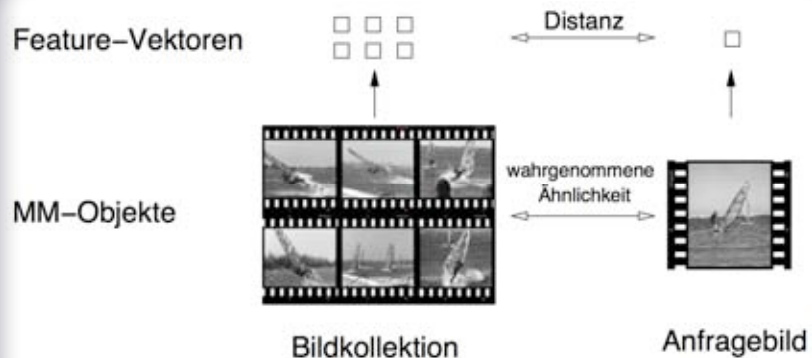


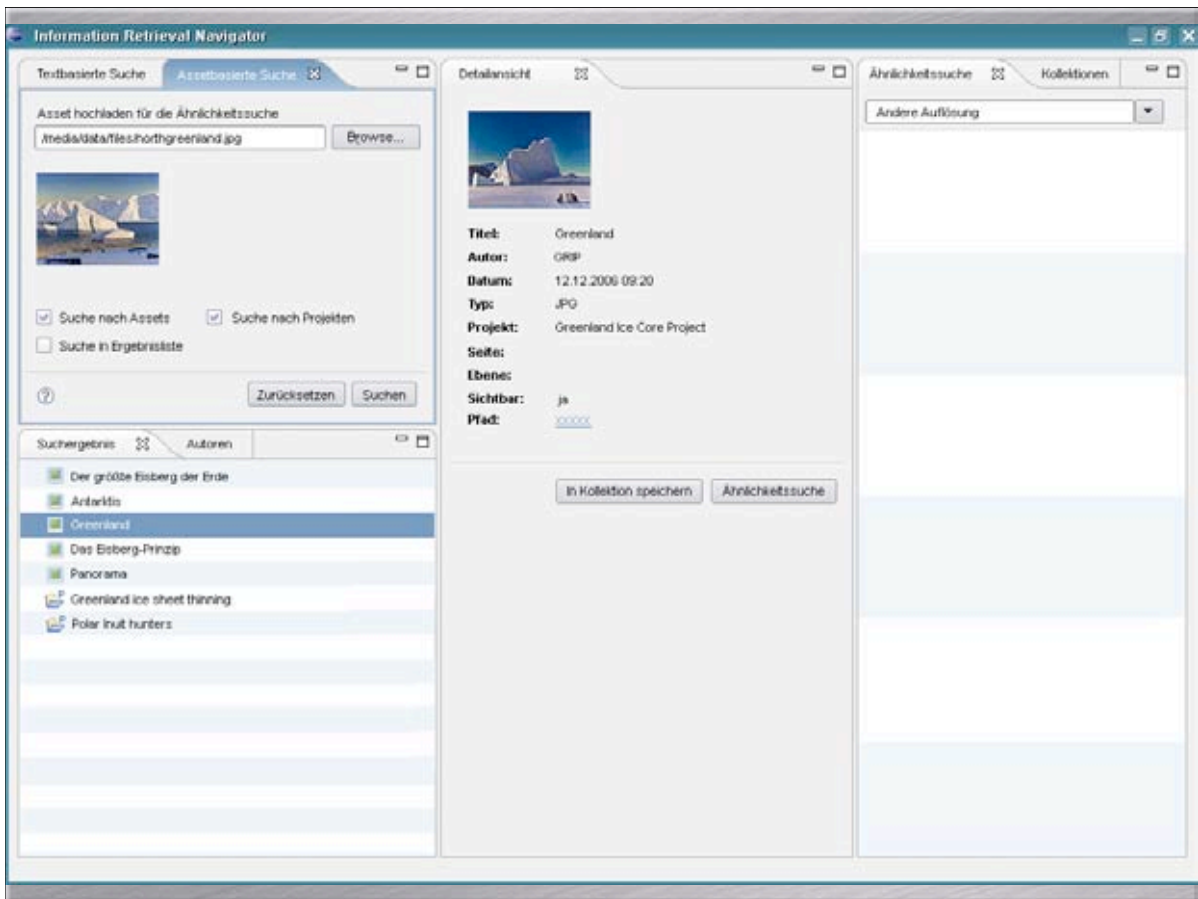
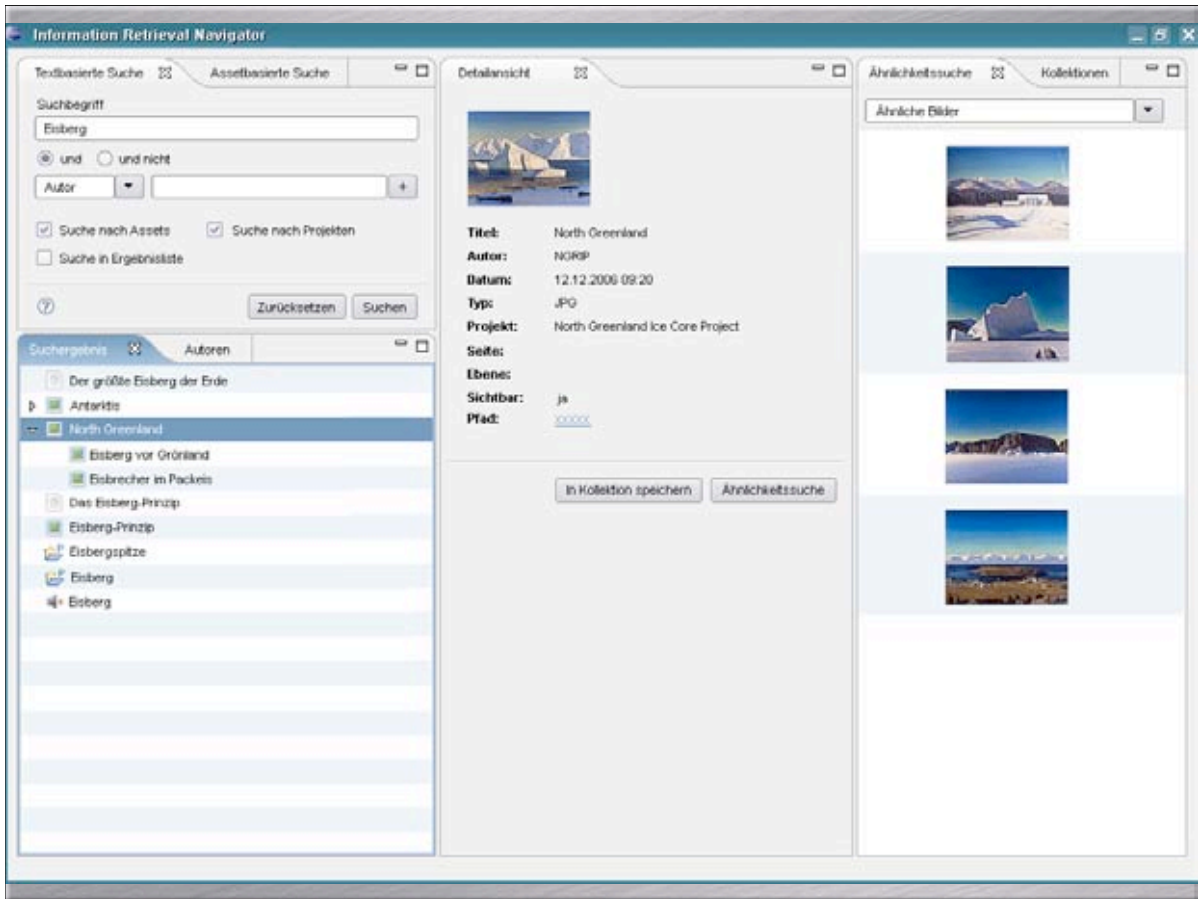
- High level feature:
Eisberg, Fisch, IRT-Gebäude,
Personen

semantische Lücke

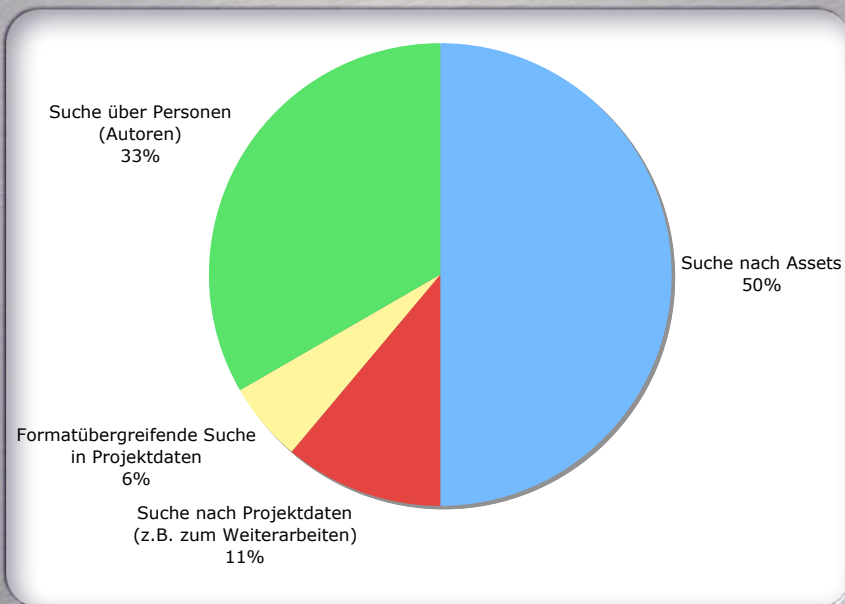
- Low level feature:
Farbverteilung, Textur, Form

CBR • Feature Extraktion





Anforderungen an das Retrieval



ILM-Studie TUB/FHTW Berlin Stand 2/2007 Interimsergebnisse (to be published)

The screenshot displays the 'Information Retrieval Navigator' interface. The search term 'Eisberg' is entered in the search field. The search results are displayed in a list on the left, including 'Der größte Eisberg der Erde', 'Arktis', 'North Greenland', 'Eisberg vor Grönland', 'Eisbrecher im Packeis', 'Das Eisberg-Prinzip', 'Eisberg-Prinzip', 'Eisbergspitze', 'Eisberg', and 'Eisberg'. The central part of the interface shows a network diagram with nodes representing authors: Berta (9), Klaus (14), Susanne (28), and Uwe (6). The right side of the interface shows a list of authors: Berta, Klaus, Susanne, and Uwe, with a sub-list of works under Berta: Eisberg, Eisberg vor Grönland, Eisbrecher im Packeis, Eisberg, and Eisberg.

PREVIEW

Prozessautomatisierung durch Generische Content Transformation. (GCT)

Vielen Dank für Ihr Interesse!

Michael A. Herzog
FHTW Berlin/TU Berlin
herzog@fhtw-berlin.de

IRT Symposium IPTV • 13. März 2007