mediaclipping.de

Automatische Metadaten Generierung im IPTV

@Forum IPTV: Technologie und Plattformen

Nils T. Kohle, mediaclipping GmbH

Fragestellung

Wie kann ich aus (User Generated) Rich-Media Content automatisch Metadaten generieren?

- Zur Verbesserung der Archivierung (nach Stichworten “Tagging”, etc.)
- Zum einfacheren Auffinden von Inhalten (es kann nur das gefunden werden, was “jemand” eingegeben hat)
- Zur Verknüpfung mit weiteren Informationen (Ads, Links, etc.)

Agenda

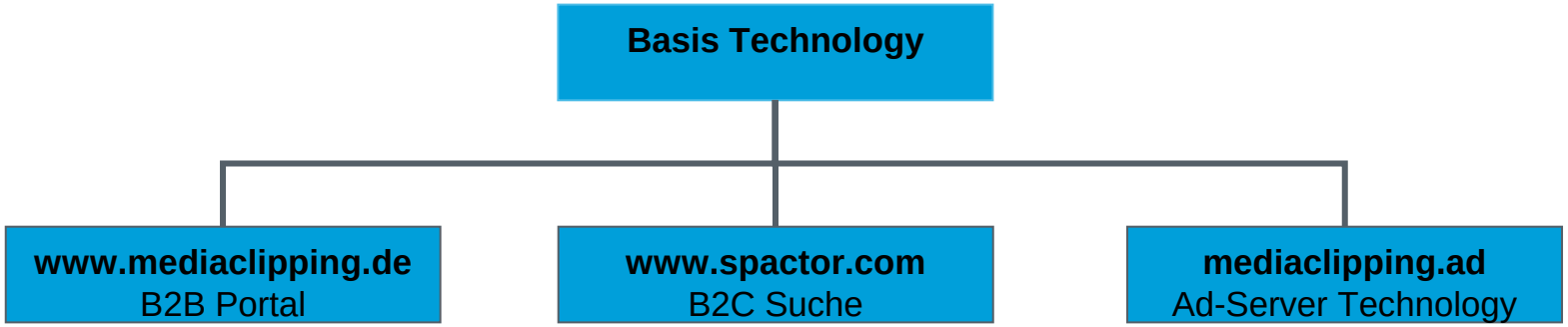
1. mediaclipping.de
2. Technologie
3. Business Model
4. Fazit

1. mediaclipping.de – Unternehmensportrait

- Start der Entwicklung Mitte 2005
- Gründung 1. November 2006 in Bremen
- Gesellschafter:
  - Management,
  - Redalpine Capital
  - nwk nordwest Kapitalbeteiligungsgesellschaft der Sparkasse Bremen mbH,
- Mitarbeiter: 16
- Ausgezeichnet:
  - 1. Platz Landeswettbewerb Deutscher Gründerpreis
  - Gewinner Deutscher IPTV Award 2007, Kategorie „Sonderpreis“
  - enable.de / Financial Times (Finalist 2007)

Standbeine – Geschäftsfelder auf Technologie-Basis

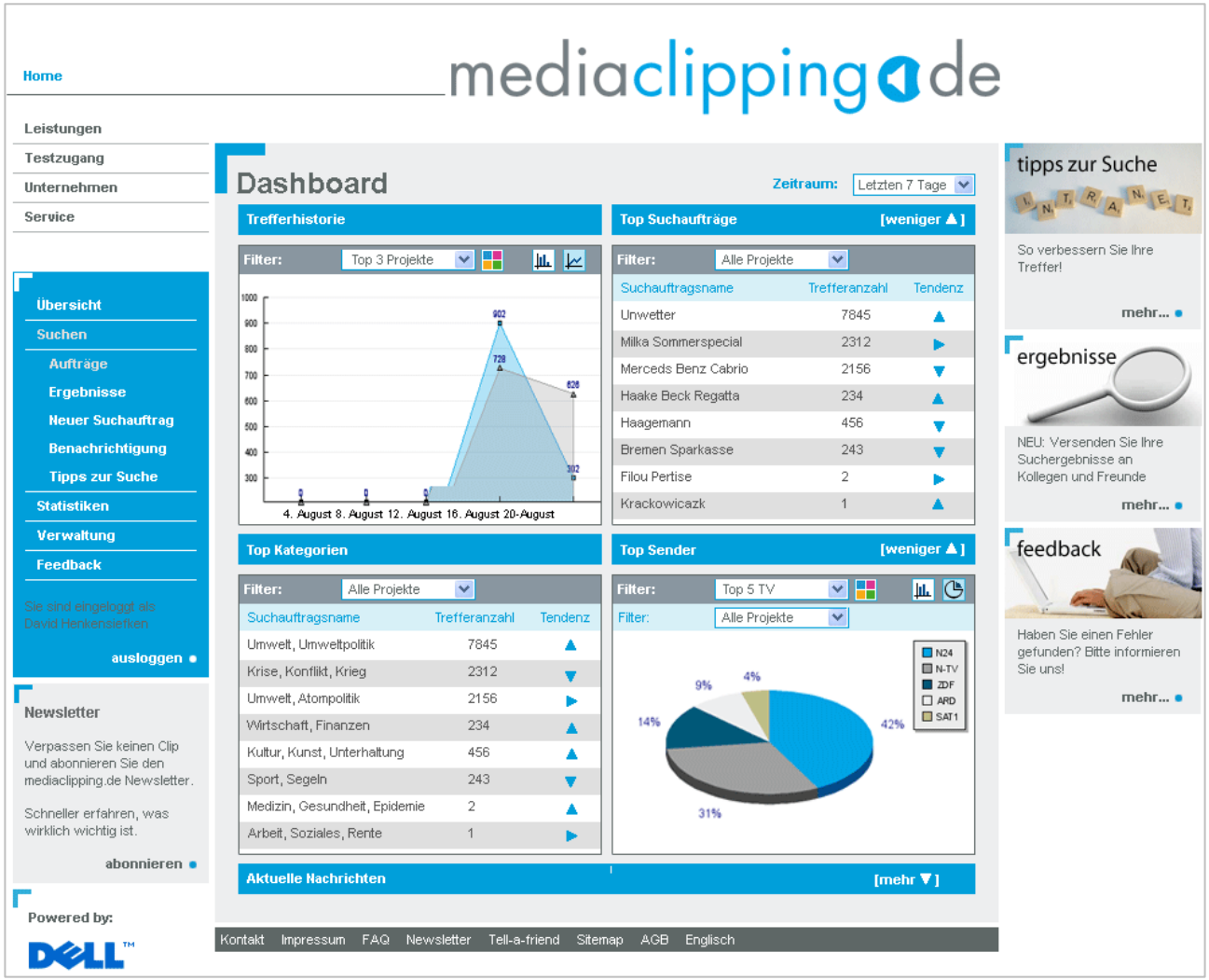
- Drei Geschäftsfelder, basierend auf der mediaclipping Basis-Technologie:
  - mediaclipping.de Service-Portal für das vollautomatisierte Clipping nach beliebigen Begriffen wie Marken, Namen, etc. in IP-basierten Audio-/ Video-Streams für B2B und B2C, z.B. Video Suche
  - B2C Suche spactor.com (live media search) und damit verbundene White-Label Lösungen auf der vorhandenen Technologie zur Aufarbeitung (Indexierung, Verschlagwortung) von IP-basierten Daten
  - Ad-Server zur kontext-sensitiven Auslieferung von Adwords-ähnlichen Anzeigen in Streaming Medien in Echtzeit (in der Entwicklung)



B2B Portal Screenshot

mediaclipping.de

Screenshot Dashboard



Unternehmen – Referenzen B2B

2. Technologie

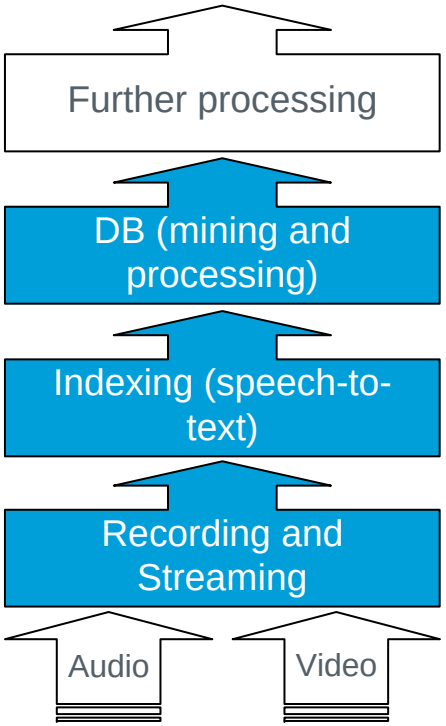
Wieso wir auf das Thema Metadaten Generierung gekommen sind?

Basierend auf unserer Technologie erschaffen wir vollautomatisch von Suchmaschinen einfach lesbare Inhalte:

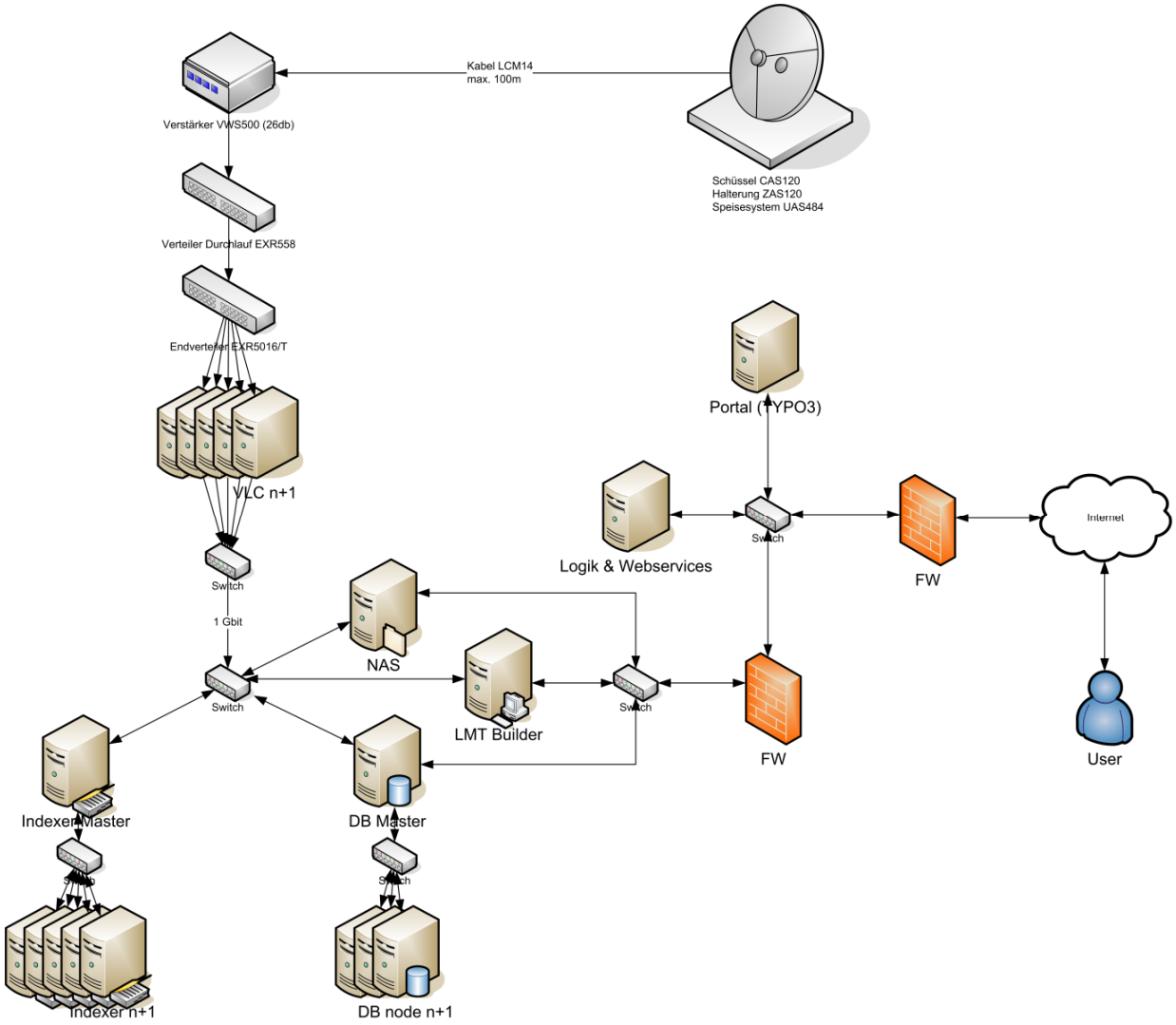
TEXT.

Basis-Technologie Echtzeit Indexierung von Rich-Media Streams

- Die Basis Technologie ermöglicht die Indexierung (speech-to-text) von TV, Hörfunk, Video und Podcast mittels einer vollautomatisierten und sprecher-unabhängigen Spracherkennung und die vollautomatische Bereitstellung von gefilterten Ergebnissen.
- Pro Eingabekanal (Radio, TV, Video, Podcast) wird ein Indexer benötigt, dieser muss mehrere Dinge tun:
  - Aufnehmen der Daten
  - Suchen der Daten im Wortschatz
  - Speichern der Daten in der Datenbank
  - Ablegen der aufgenommenen Daten in komprimierter Version (Realmedia, Windows Media, mp3)



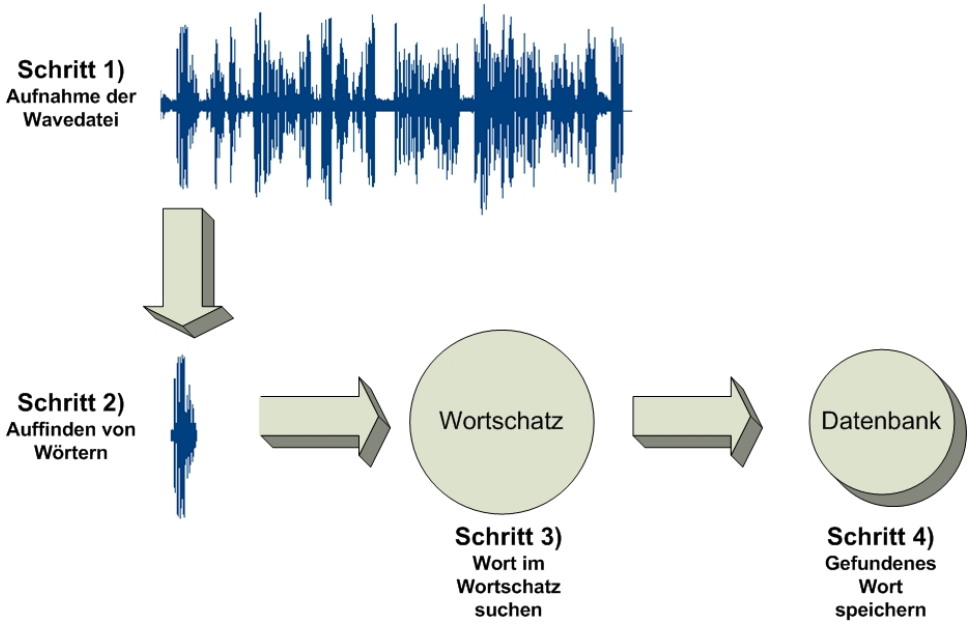
Technische Skizzierung Darstellung des Prozessablaufs



mediaclipping de

Sprachtechnologie Sprecherunabhängige Indexierung in Echtzeit

- Einsatz sprecherunabhängiger Sprachalgorithmen
  - Erkennungsquote von Wörtern in den gescannten Audiodaten (abhängig von der Datenbasis) mind. 80-85%
  - Außenaufnahmen mit Hintergrundgeräuschen/ Dialekte ca. 35%
- Bei sinnentfremdeter Textwiedergabe Verifikation möglich
- Laufende automatisierte Wortschatzerweiterung durch LMT-Modell



mediaclipping de

3. Business Model

Es gibt zur Zeit kaum Refinanzierungsmodelle für IPTV Anbieter, da die “stumpfe” Einblendung von Werbemitteln ins Web 1.0 gehört.

Für kontext-sensitive Werbung fehlen bislang die Verknüpfungen, diese schaffen wir mit der Metadaten Generierung.

Die Zukunft bei der Aussteuerung von Ads im Rahmen von Rich-Media sieht mehrere Möglichkeiten:

- Behavioural Ad-Serving: Gewohnheiten des Users über immer wieder gewählte Themen/Begriffe werden kontrolliert und dokumentiert, auf der Basis der gesammelten Informationen wird das passende Banner zum Profil eingespielt
- Kontext-sensitive Aussteuerung von Ads ist technisch deutlich komplexer (Sprach-Technologie) und kann – z.B. in Verbindung mit der Kategorisierung zu einem erweiterten Google-AdWords werden
- Tags verknüpfen (durch den User generiert) Audio/Video Inhalte mit bestimmten Themen
- Captions, in den USA verpflichtend für Audio/Video Inhalte können genutzt werden um Werbung passend zum Inhalt anzuzeigen

Wir gehen davon aus, dass es in Zukunft alle Formen geben wird um eine möglichst hohe Refinanzierung von Content zu ermöglichen.

Kontext-sensitiv vs. Behavioural Beispiel für behavioural Ad-Aussteuerung

- Keine Informationen über den Videoclip. Stattdessen werden Daten über den Benutzer gesammelt (Datamining) und Profile über sein Verhalten erstellt
- Bsp.: Der Kunde war auf eBay.de. Danach wechselt er zu Spiegel.de. Dieses wird anhand von Tracking-Cookies aufgezeichnet. Der Ad-Server steuert nun die entsprechende Werbung aus: In diesem Falle für ein Angebot auf Ebay.de.
- Nachteil: Der Aktuelle Videoclip auf Spiegel.de handelt über eine Feuerkatastrophe/Brand. Die Werbung hat den Slogan „Heiß macht reich“ (reales Beispiel, 25.06.2007. 13:35 Uhr)

mediaclipping de

Kontext-sensitiv vs. Behavioural Beispiel für kontext-sensitive Ad-Aussteuerung

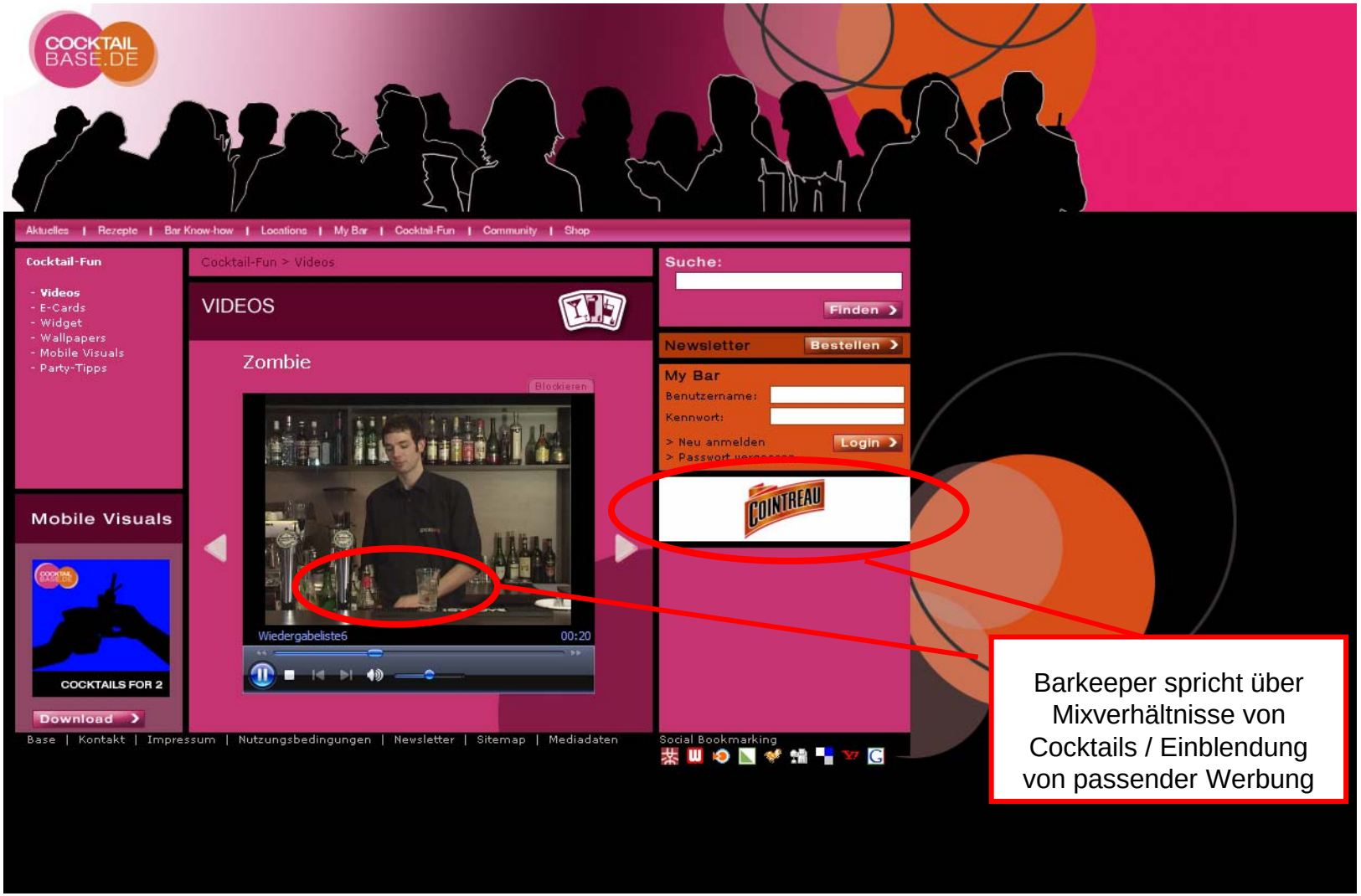
- Informationen über den Videoclip sind vorhanden. Keine Informationen über den/die Benutzer
- Bsp.: Der Aktuelle Videoclip auf Spiegel.de handelt über eine Feuerkatastrophe/Brand. Die Werbung ist passend (kontext-sensitiv) zum Inhalt des Videos ausgesteuert, z.B. Feuerschutzversicherungen.
- Vorteil: Die Werbung steht im Bezug zum Inhalt des Videoclips und somit im Gesamtkontext. Es ist nicht nötig zu wissen, wer exakt der Rezipient ist.

Überleitung Metadaten-Generierung

Wenn wir es schaffen, aus den Rich-Media Inhalten Metadaten zu generieren, kann es wie folgt aussehen.

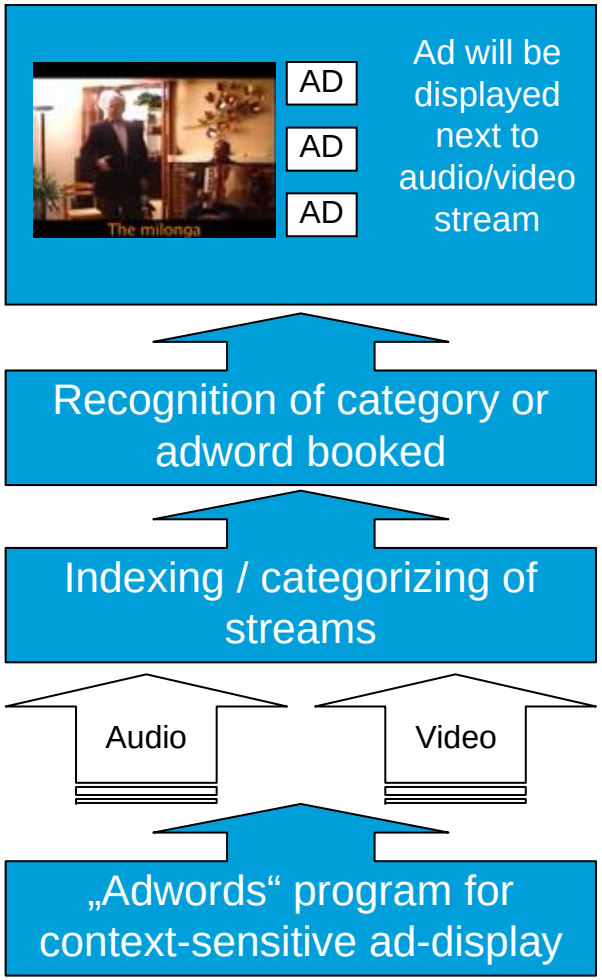
Suche mit Hilfe von Metadaten Beispiel (I/II)

Suche mit Hilfe von Metadaten Beispiel (II/II)



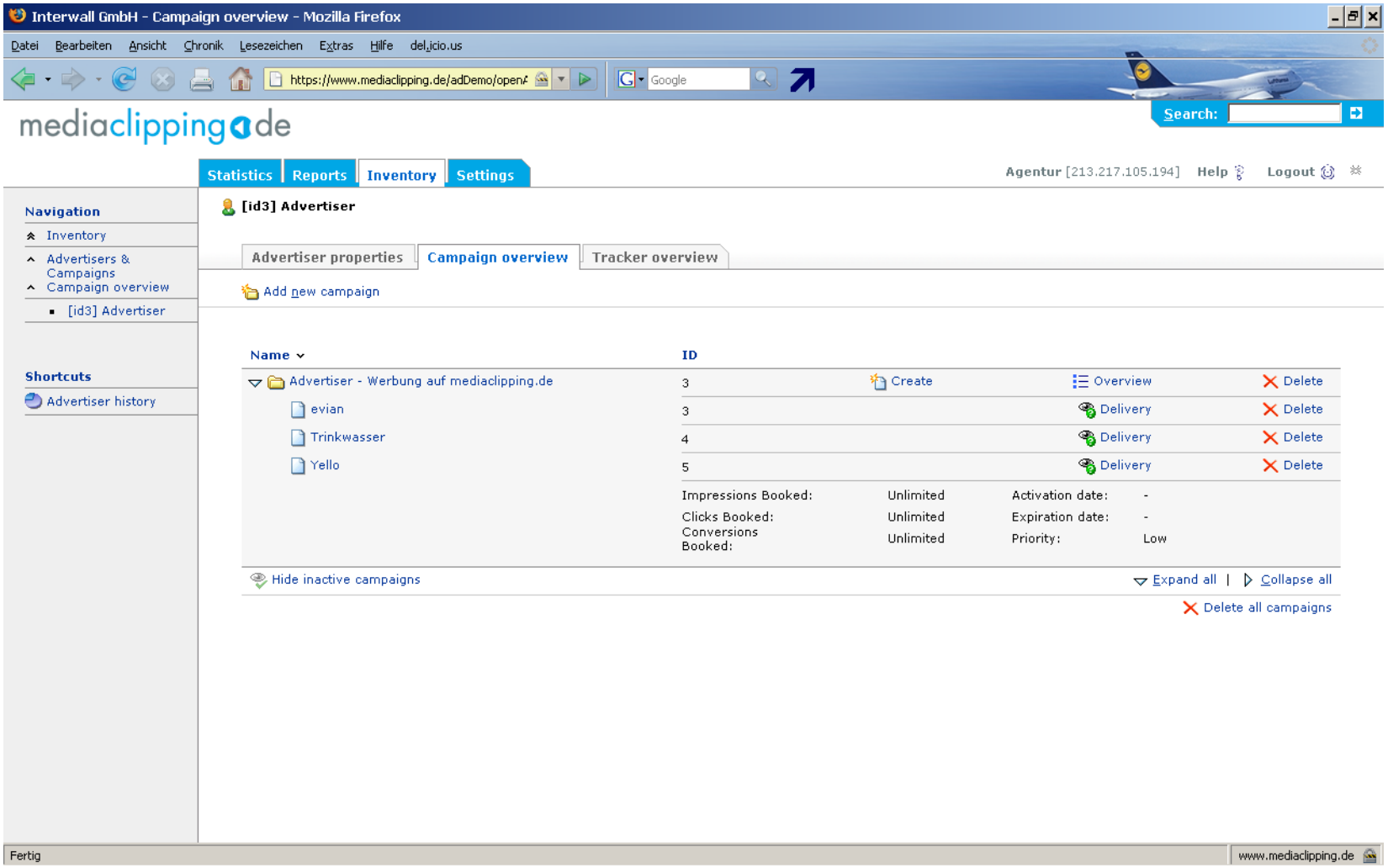
Metadaten Generierung Erläuterung des kontext-sensitiven Ad-Servers

- Der mediaclipping Ad-Server basiert auf der Basis-Technologie des Unternehmens und ermöglicht
  - Echtzeit Erkennung von Worten in IP-basierten Audio-/ Video-Streams, d.h. die Generierung von Metadaten
  - Die Darstellung von Ads nach Nennung von Wörtern (z.B. Bier) oder Kategorien (Getränke)
- Der kontext-sensitive Ad-Server ist bereits im ersten (Test-) Einsatz
- Ziel des Ad-Servers ist es, Anbietern von Rich-Media Content eine weitere Refinanzierungs-Möglichkeit zu bieten



mediaclipping de

Ad-Server Backend-Ansicht



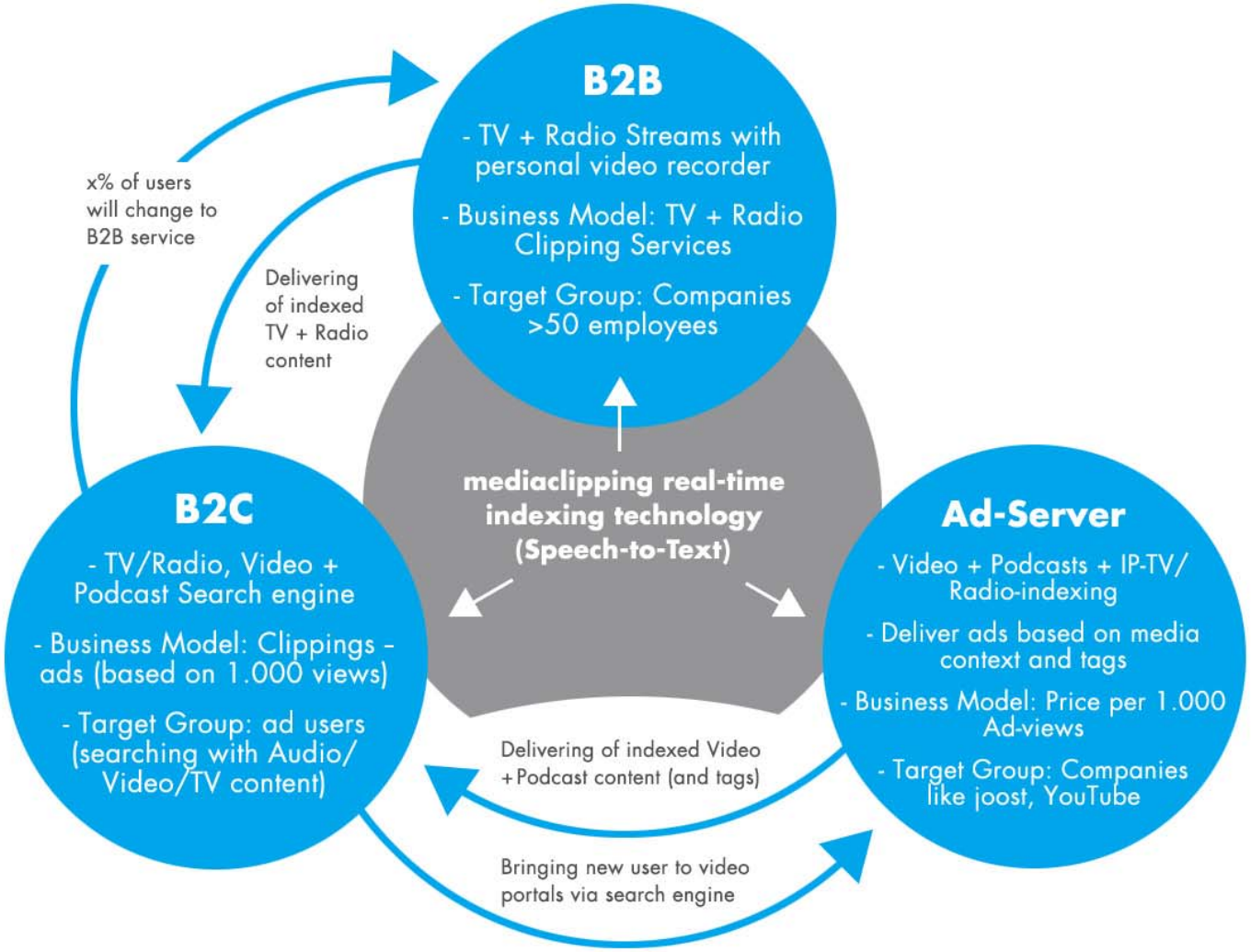
mediaclipping.de

Ad-Server Backend-Ansicht

Ad-Server Beispielhafte Integration

mediaclipping de

Business Modell B2B + Ad-Server + B2C



4. Fazit

- Geschäftsmodelle werden in Zukunft im IPTV Bereich vielfach durch Werbung finanziert werden
- Werbung muss in Zukunft allen Prognosen nach möglichst gut zum Content / User passen, das führt zur Nutzung von kontext-sensitiven Systemen der Aussteuerung
- Die Generierung von Metadaten ermöglicht das Aufsetzen neuer Geschäftsmodelle zur Refinanzierung von Services, unabhängig vom User-Profil

Kontakt für weitere Informationen

Nils T. Kohle – Geschäftsführer

mediaclipping GmbH

Konsul-Smidt-Straße 8r / Port 04

28217 Bremen

Telefon +49 (0) 421 333055-300

Fax +49 (0) 421 333055-333

E-Mail n.kohle@mediaclipping.de

Internet www.mediaclipping.de / www.spactor.com