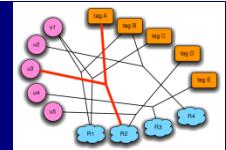


Lehrstuhl für Informations-
und Kommunikationsdienste



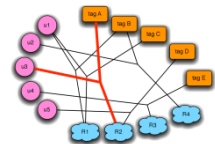
Konzeption eines Frameworks für die Evaluation von Tag-Suggestion-Algorithmen

Martin Garbe
Steffen Oldenburg
Lukas Zielinski
Prof. Dr. Clemens Cap
(Universität Rostock)

08.05.2008 (Bamberg)

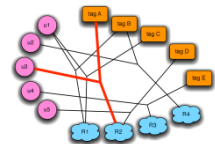
Übersicht

- **Tags und Folksonomien**
- **Tag-Suggestions**
 - Definition
 - Generierung
 - Vergleich der Ergebnisse
- **Konzeption des Evaluationsframeworks**
- **Messungen am Prototyp**
- **Zusammenfassung**



Information Retrieval und Tagging

- **Indexierung: Automatisch vs. Manuell**
 - Volltextsuchmaschine
 - Klassifizierung
 - Ontologien
 - *Tagging-Systeme (del.icio.us, flickr, youtube)*
- **Trend geht zu benutzergeneriertem Content (Web 2.0)**



Tags, Folksonomien



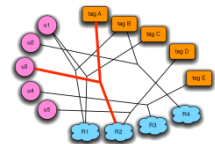
Tags

- Water
- Lakes
- Rivers
- Creeks
- See
- Bergsee
- Bach
- Wildstrubel Berge Schweiz
- pyrenäische Berghunde
- Alpen
- Bergtouren
- Bergwandern
- pyreanean mountain dogs
- mountains
- switzerland
- hiking
- alps
- glaciers
- summits
- dogs

- Tags sind von Benutzern vergebene Wörter zur Indexierung von Ressourcen.
- Eine Folksonomie ist eine Sammlung von Tags (T), Benutzern (U), Ressourcen (R) und Tupeln ($Y = T \times U \times R$).

$$F = (T, U, R, Y)$$

- Auch $Y = T \times R$ möglich



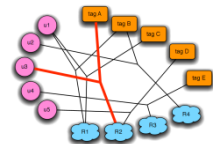
Tags, Folksonomien

■ Vorteile gegenüber hierarchischer Klassifizierung

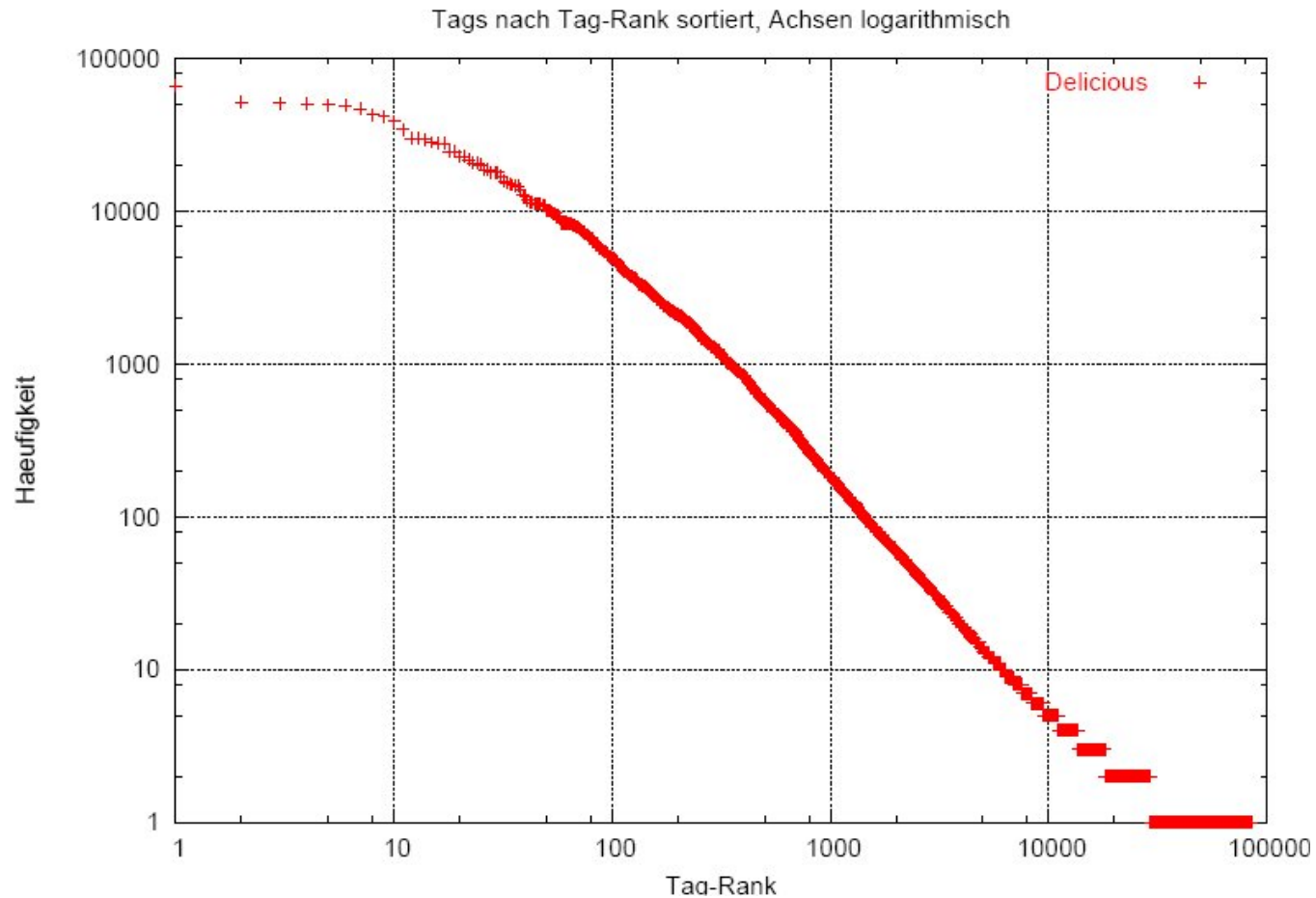
- Freie Wahl von Tags, kein Hierarchiezwang
- „Entdeckungscharacter“
- Nutzen für Allgemeinheit durch viele egoistische Ziele
- Tags gut geeignet zur Beschreiben von Binärdaten (Bilder, Videos)

■ Nachteile

- Mangelnde Präzision bei Suche, geringer Recall (syntaktische Variationen, populäre Tags, keine Hierarchie)
- Organisatorische Tags
- Spam

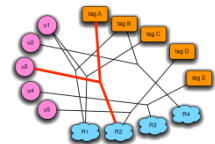
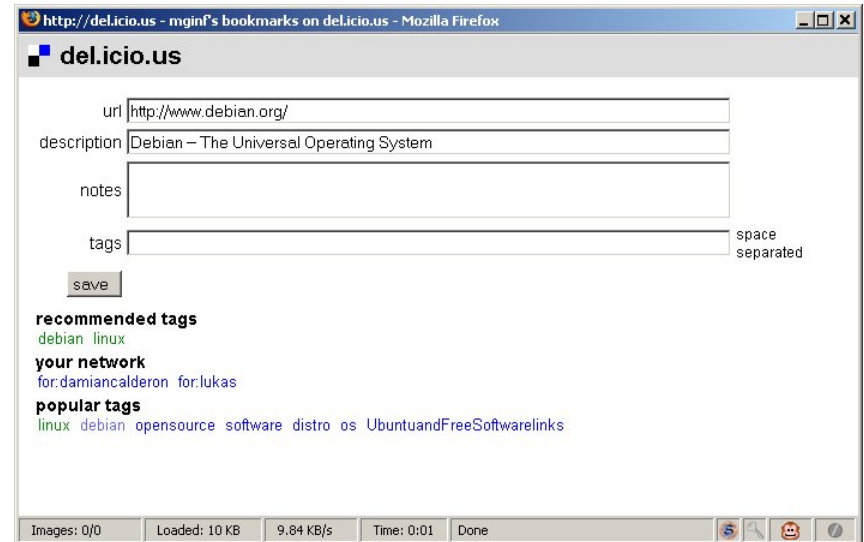


Power law



Tag-Suggestion

- Vorschlagen von Tags durch das System
- Erleichtert Tagging
- Ausbildung eines etablierten Vokabulars
- Höherer Recall
- Senkt Selektivität von Tags



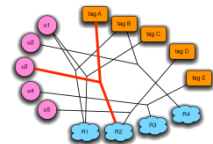
Generieren von Tag-Suggestions

■ Erzeuge Tags aus

- bestehenden Tags der Folksonomy
- Inhalt einer Ressource

■ 4 Verfahren wurden untersucht:

- Tag-Vektoren
- Co-Occurrence Matrix Clustering
- Scoring-Verfahren
- Inhaltsanalyse mit TF-IDF



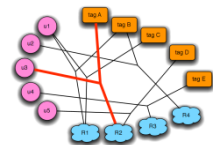
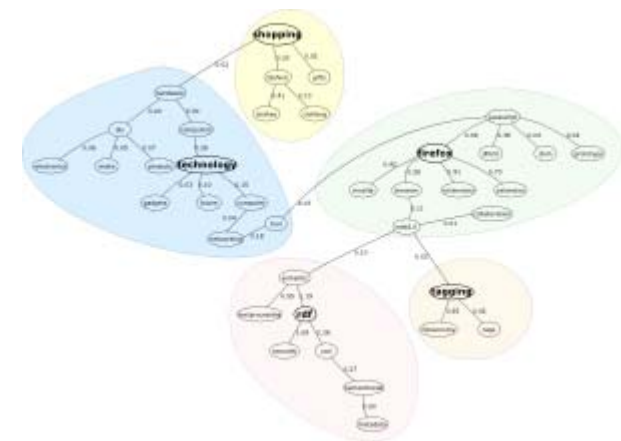
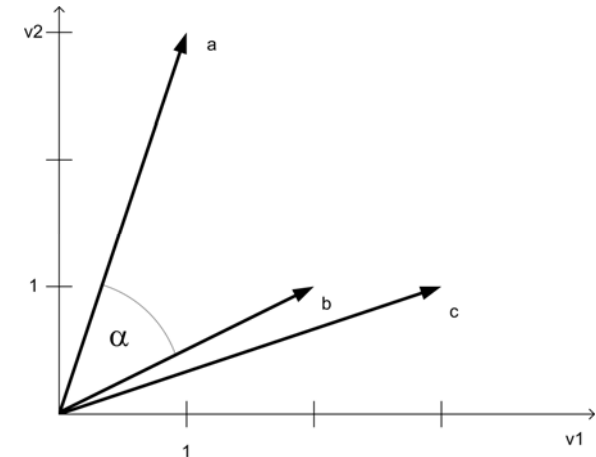
Generieren von Tag-Suggestions

■ Tag-Vektoren

- Tag durch Vektor repräsentiert
- Ressourcen stellen Komponenten des Vektors dar
- Ähnlichkeit von Tags => Cosinus

■ Co-Occurrence Matrix Clustering

- Gemeinsames Auftreten von Tags => Co-Occurrence
- Co-Occurrence-Grad zw. Tags bestimmen
- Semantische Ähnlichkeit von Tags durch Clustering finden
- K-Metis anstatt Spektralem Clustering nutzen



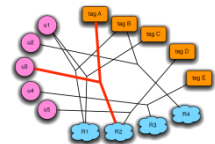
Generieren von Tag-Suggestions

■ Scoring-Verfahren

- Grundlegender Gedanke: Gute Tags genügen gewissen Kriterien z.B. Coverage, Score, $P(t_i | t_j)$
- Belohne Tags, wenn von vielen genutzt
- Bestrafe Tags, wenn sie große Schnittmenge an Ressourcen mit anderen Tags annotieren

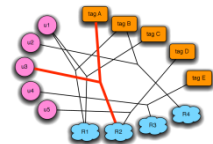
■ Inhaltsanalyse mit TF-IDF

- Extrahiere Keywords aus Inhalt der Ressource
- Berechne TF-IDF-Wert



Vergleich von Tag-Suggestions

- **Wie gut oder schlecht ist ein Tag-Suggestion-Algorithmus?**
- **Ausgabe (vorgeschlagene Tags) bewerten**
- **Methoden:**
 - Tag-Frequenz
 - Tag-Coverage
 - Tag-Overlap
 - ...



Vergleich von Tag-Suggestions

■ Tag-Frequenz

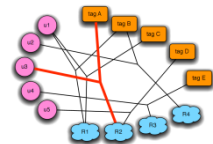
- Anzahl der Vorkommen eines Tags

■ Tag-Coverage

- Wie Tag-Frequenz, jedoch ohne Dopplungen
- Anzahl der verschiedenen Ressourcen mit Tag X

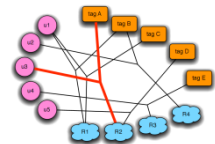
■ Tag-Overlap

- Wie unterschiedlich sind mit Tag X,Y versehene Ressourcen
- Beschreiben Tags ähnliche Konzepte?

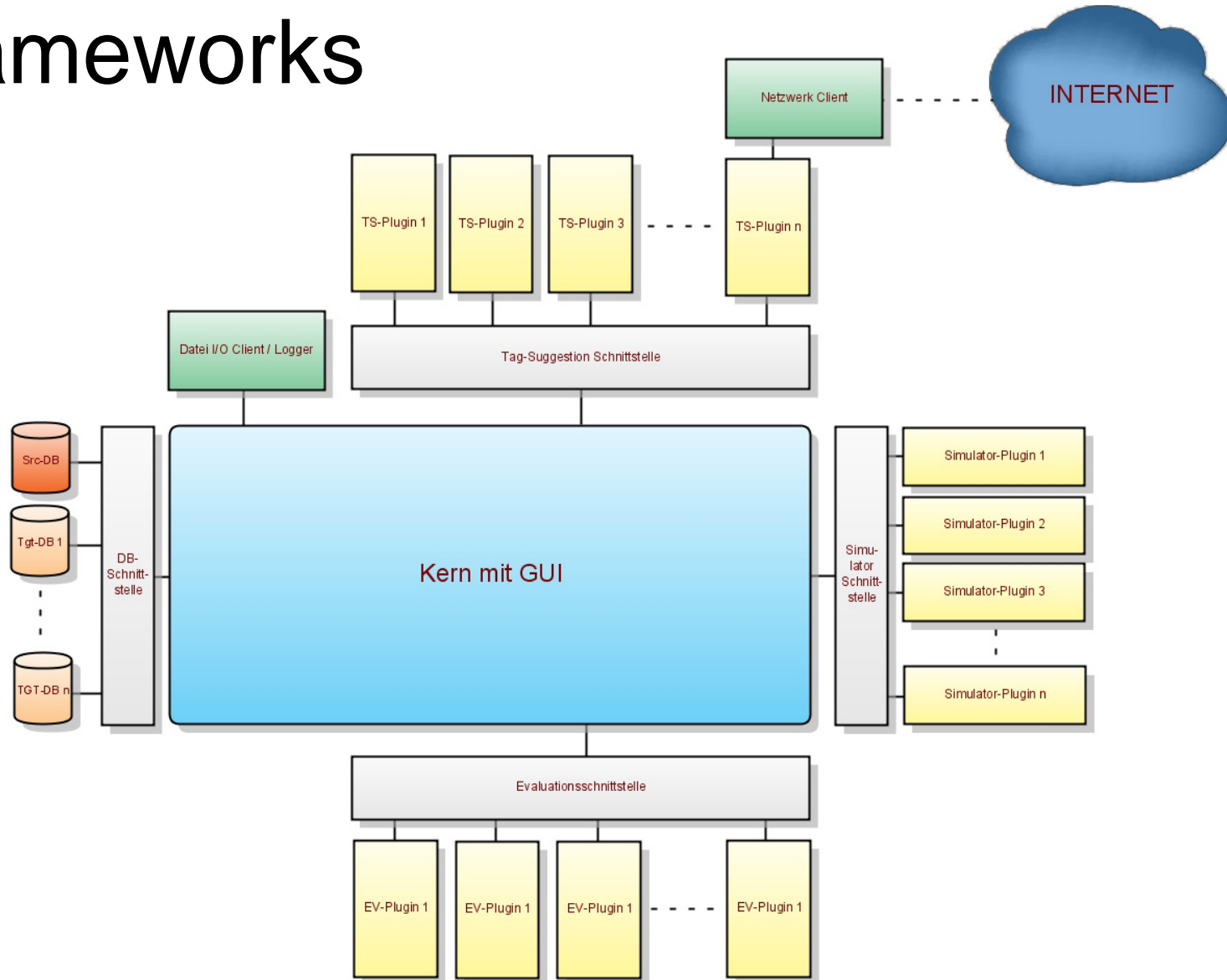


Konzeption des Tag-Evaluation-Frameworks

- **Bisherige Evaluationen auf unterschiedlicher Datenbasis nach verschiedenen Kriterien**
- **Anforderungen:**
 - Architektur zur Entwicklung und Auswertung von Tag-Suggestion-Algorithmen
 - Unterstützung von Broad Folksonomies
 - Mehrere Quell-Folksonomien unterstützen
 - Leichte Erweiterbarkeit
 - Tag-Suggestion-Algorithmen
 - Meßgrößen
 - Simulatoren

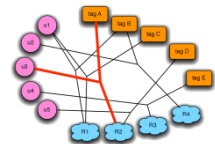


Konzeption des Tag-Evaluation-Frameworks

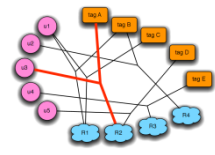
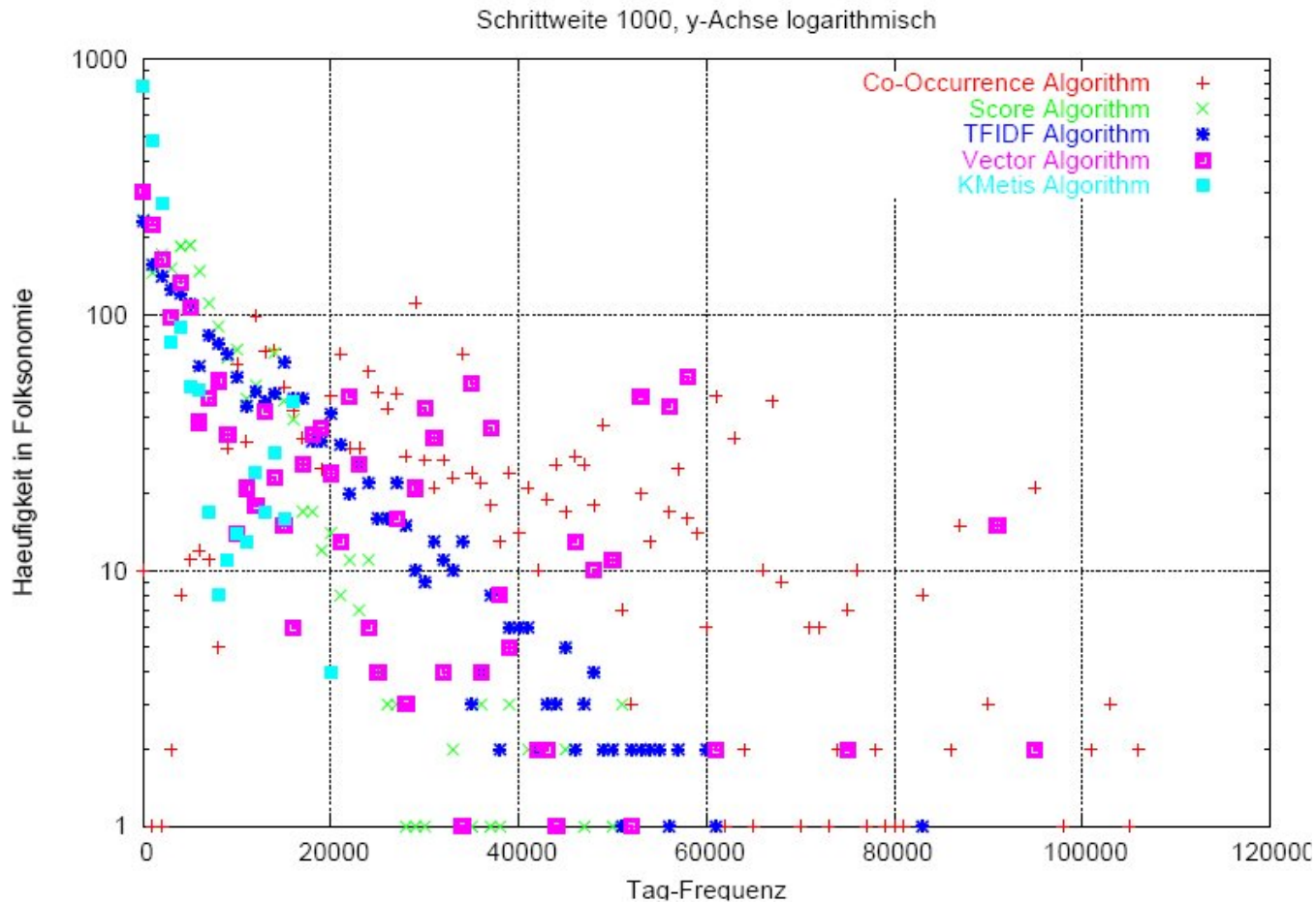


Messungen

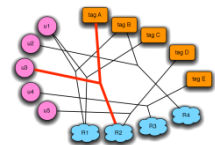
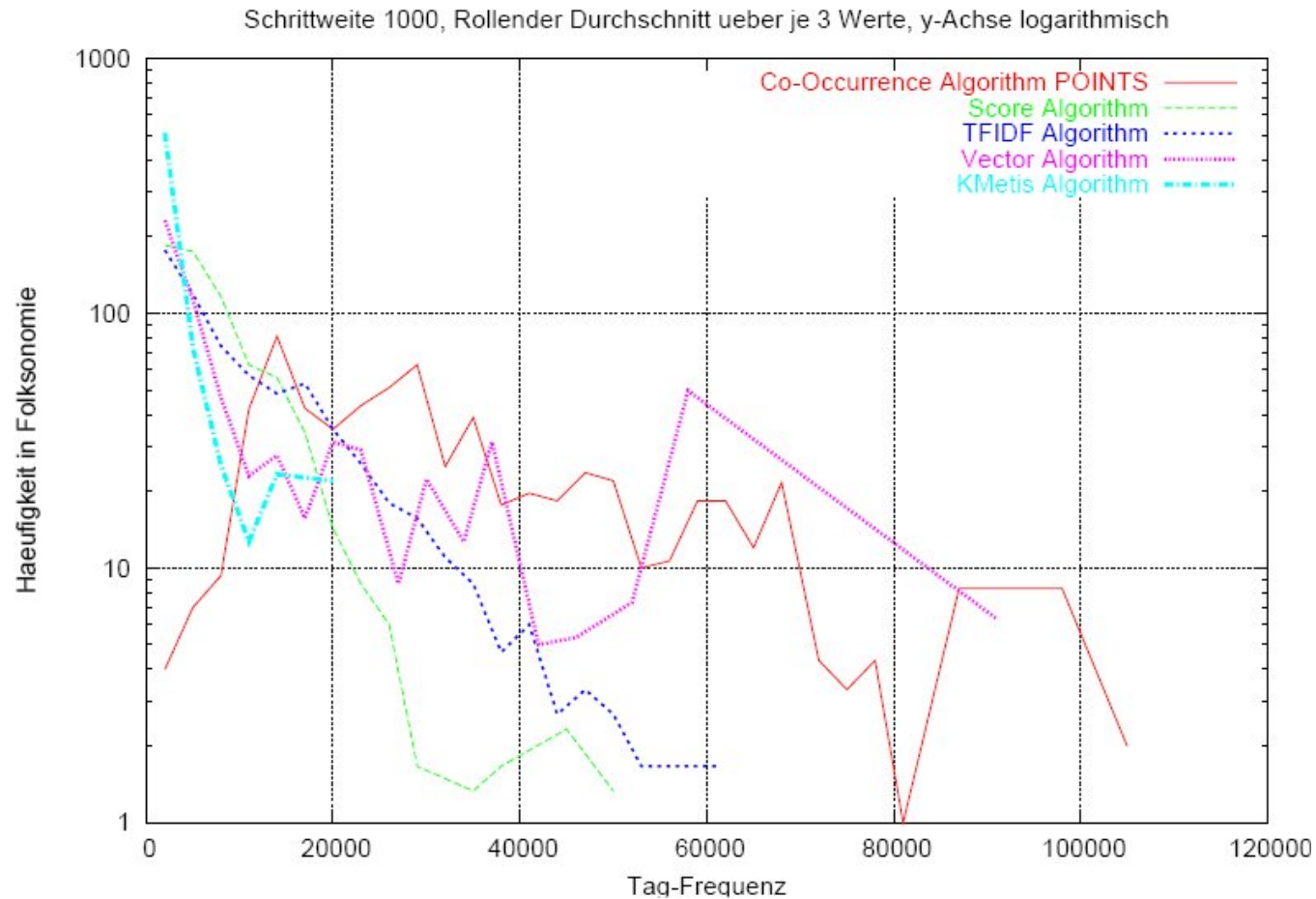
- Simulator startet Algorithmus 2000 mal und erwartet maximal 10 Tag-Suggestions
- Aus Performancegründen 80.000 statt 8.000.000 Tag-Assignments aus Quell-Folksonomy
- Vorgeschlagene Tags in neuer Folksonomy speichern



Messungen: Tag-Frequenz

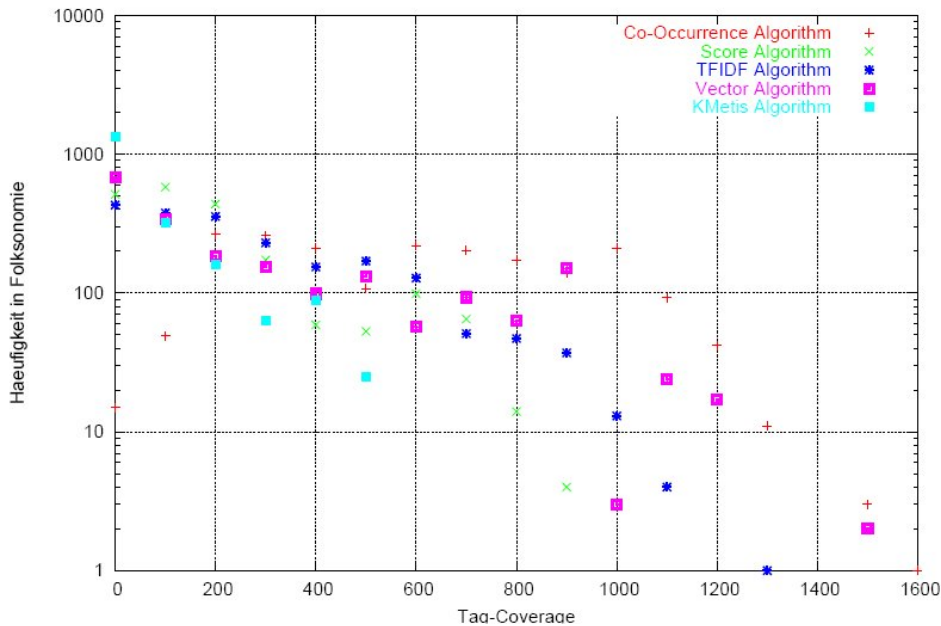


Messungen: Tag-Frequenz

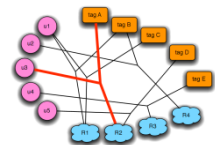
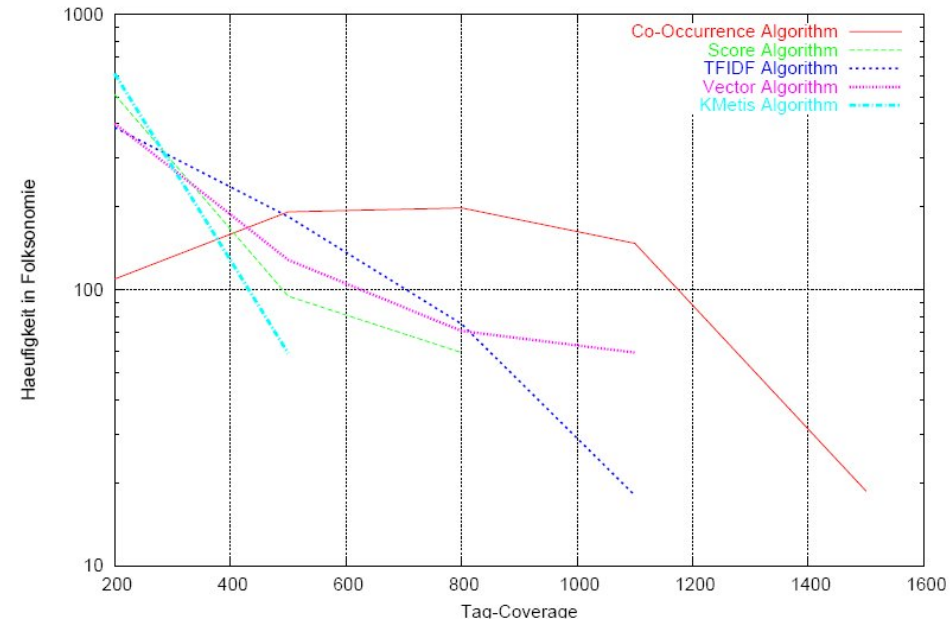


Messungen: Tag-Coverage

Schrittweite 100, y-Achse logarithmisch

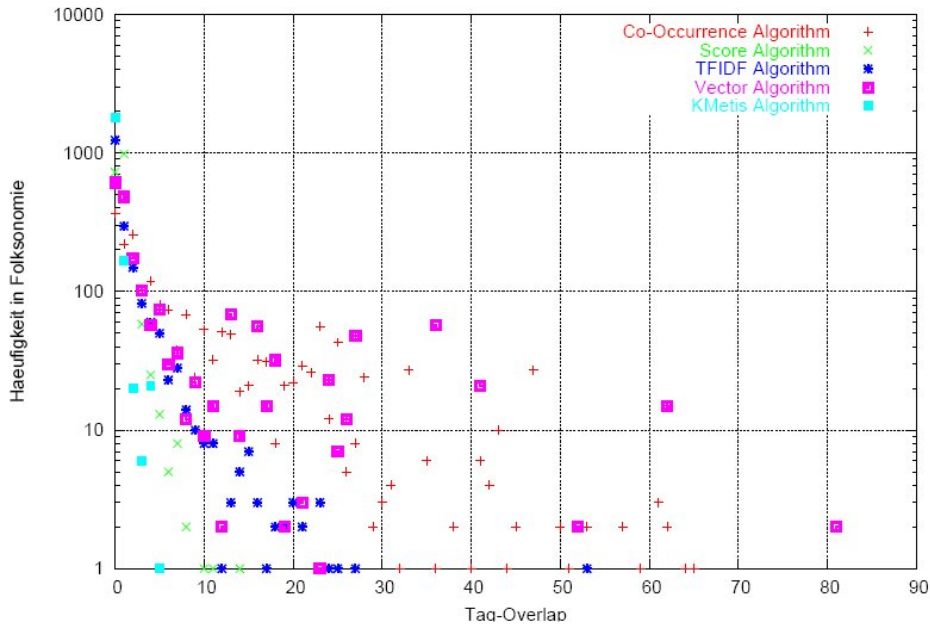


Schrittweite 100, Rollender Durchschnitt ueber je 3 Werte, y-Achse logarithmisch

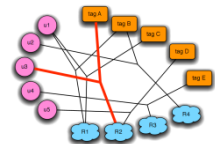
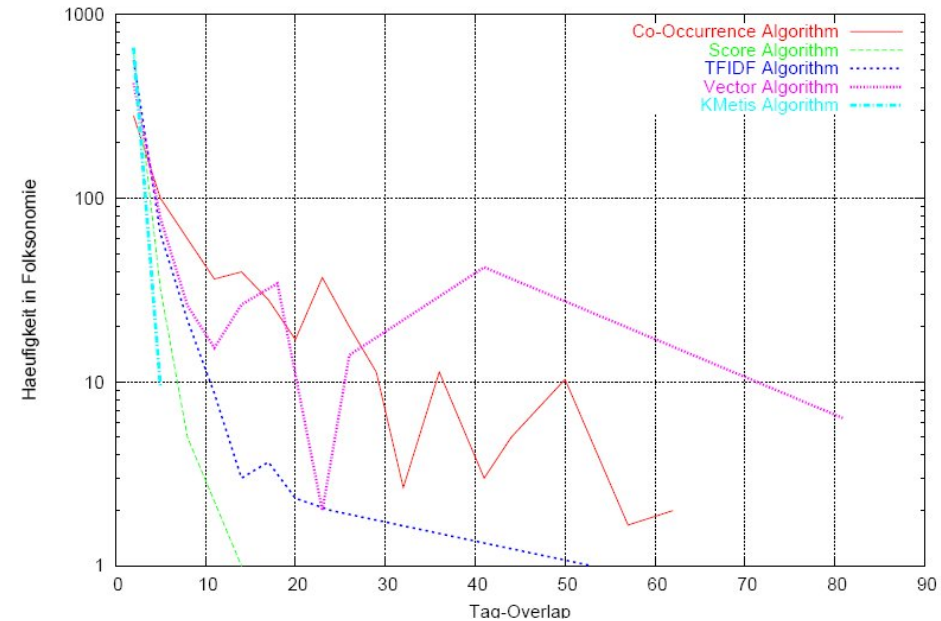


Messungen: Overlap

Schrittweite 1, y-Achse logarithmisch

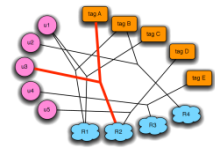


Schrittweite 1, Rollender Durchschnitt ueber je 3 Werte, y-Achse logarithmisch



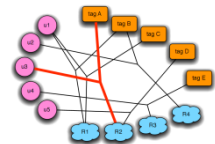
Zusammenfassung

- **Tag-Suggestion-Algorithmen lassen sich quantitativ vergleichen**
- **Je nachdem welche Ziel verfolgt wird, haben die Tags unterschiedliche Güte (z.B. allgemeine Tags, präzise Tags)**
- **Framework erlaubt einheitliches Testen und Auswerten von Tag-Suggestion-Algorithmen**
- **Kombination von Tag-Suggestion-Algorithmen nutzen**



Ausblick

- **Verbesserung der Performance des Frameworks**
- **Testen neuer Meßgrößen**
- **Kombination verschiedener Folksonomien als Grundlage für Tag-Suggestion-Algorithmen**



Vielen Dank für die Aufmerksamkeit.

Fragen?

