

Automatisierung der Sacherschließung mit Semantic Web Technologie

Kommunikations-, Informations-,
Medienzentrum (KIM)
IT- und Bibliotheksdienste
Universität Konstanz

SIS-GND-Matching-Automat (SIGMaMat 2014)

Verknüpfen Unbearbeitet Bearbeitet Zurückgestellt Gelöscht Suchen Statistik

Fachbereich: frz (1954) Angemeldet als [Ralph Hafner](#) > [Abmelden](#) | [Hilfe](#)

Neue Einträge										
SIS Schlagwort	Formal- stelle	Zeit- schlag- wort	Sach- begriff	Geogra- fikum	Werk	Veran- staltung	Organi- sation	Person (individua- liert)	Familie	Quelle
Apokryphen Q	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>					
A quatre pas du soleil Q	<input type="checkbox"/>	<input type="checkbox"/>								
arriver Q	<input type="checkbox"/>	<input type="checkbox"/>				<input type="checkbox"/>				
Artamene Q	<input type="checkbox"/>	<input type="checkbox"/>				<input type="checkbox"/>				
Art d'amours Q	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
Arthur Q	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		

1 - 10 von 10 Einträgen

Bearbeitete Einträge										
SIS Schlagwort	Formal- stelle	Zeit- schlag- wort	Sach- begriff	Geogra- fikum	Werk	Veran- staltung	Organi- sation	Person (individua- liert)	Familie	Quelle

Ralph Hafner, Bernd Schelling

Nürnberg, 27.05.2015

Sündenfall



https://w3.kunstnet.org/23207/gross_suendenfall.jpg

Konstanzer Systematik

Stärken

- Flexibilität
- relative Hierarchien
- Erweiterbarkeit

Schwächen

- Notationslänge (gri 900:p718:ka64a/t94)
- Merkbarkeit: Phonetik: spr 86 - spr 114

Falscher Zeitpunkt für Systematik-Wechsel

Ziele

1. Arbeitserleichterung für FachreferentInnen

- durch automatisierte Sacherschließung und
- durch Unterstützung bei Sacherschließung

2. Unerschlossenes: E-Books

3. interoperable Daten schaffen

- raus aus der Isolation:
Verschränken mit anderen Ordnungssystemen
- Systematik ins Semantic Web,
Konkordanzen als Linked Data

Konzept

Der Maschine das Verstehen beibringen

- der Konstanzer Sacherschließungsdaten
- fremder Sacherschließungsdaten

Verschränken verschiedener SE-Datenbestände

- Nutzung semantischer Verfahren zur Ähnlichkeitssuche („Semantic Web“)

Ontologie und semantisches Schließen

Klassischer Ansatz

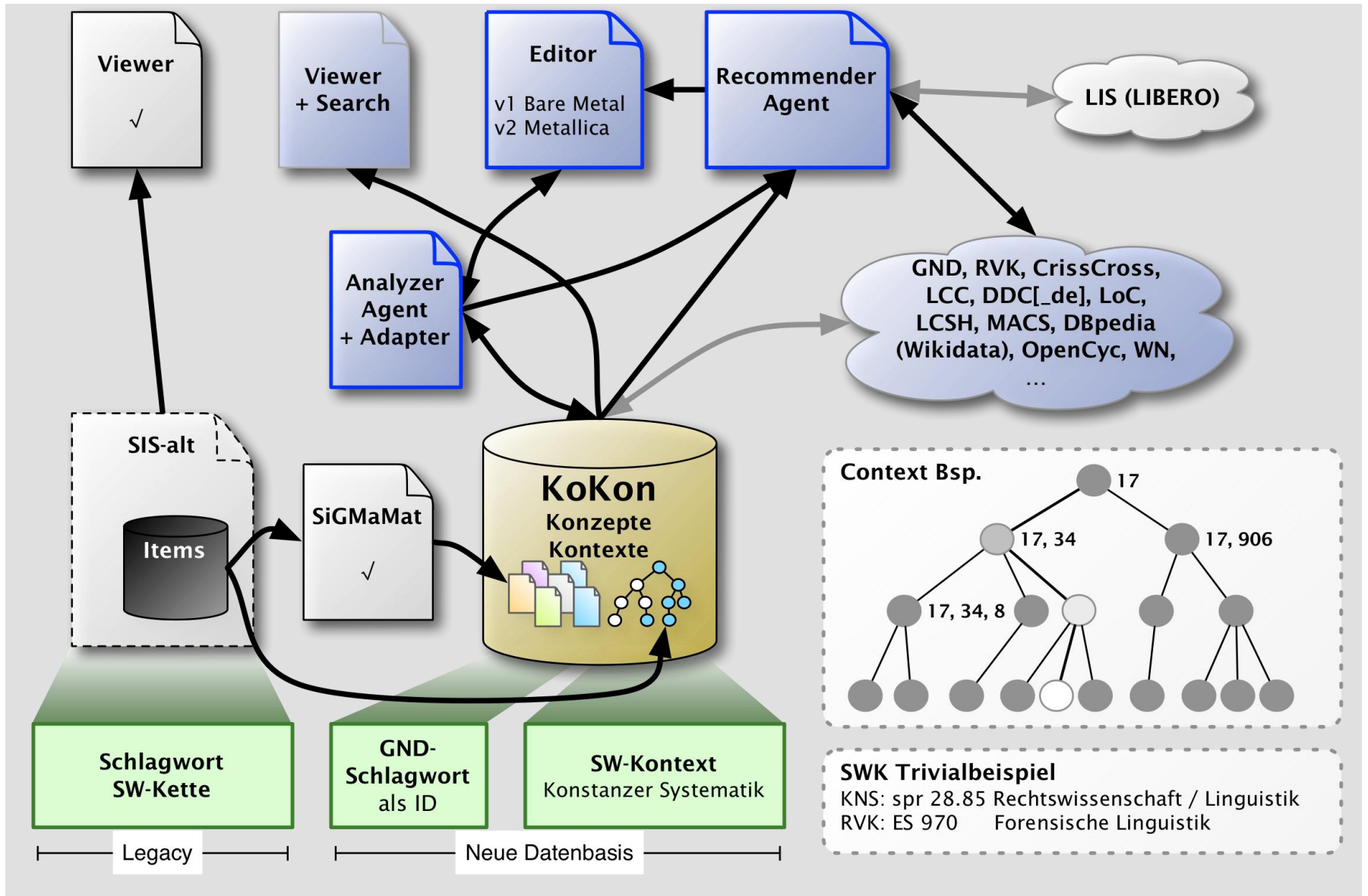
- Ontologien als „Welten“ vollständig mit einander verschalten
- Suche im gesamten Datenbestand

Konstanzer Ansatz

- von Knoten-Kandidaten zweier Systematiken ausgehen
- Unschärfe mindern durch Einbeziehung des Kontexts zur Plausibilitätsprüfung

Semantische Heuristik

- **Metrik für Distanz von SWK-KN und Knoten in anderer Systematik**
- **Metrik für Zuverlässigkeit von Aussagen (Kontext)**
- **Beginn mit „Trivialfällen“, dann Optimierung statt vollkommener Modellierung der „Welt“
→ robustes System**



Architektur

Vorbereitung zur Automatisierung

Migration der vorhandenen Sacherschließungsdaten

- GND-Identifizierer mit SiGMaMat
- Bereinigung der Strukturfehler in der Systematik
 - SIS Visualizer
 - Intellektuell durch FachreferentInnen

SiGMaMat – wozu?

Schlagwort in SIS: „Antigone“

Was ist gemeint?

- Sagengestalt
- Werk von Sophokles
- Werk von Jean Anouilh
- ...

Zeichenkette → Konzept mit ID

SiGMaMat – Funktionsweise

Vergleicht lokale SW mit SW aus Normdateien
(GND, VIAF)

Übereinstimmung → mappt

– Automappingquote: ca. 70%

Zu geringe Übereinstimmung

→ Vorschläge

→ Mapping durch FachreferentInnen



Verknüpfen

Unbearbeitet

Bearbeitet

Zurückgestellt

Gelöscht

Suchen

Statistik

Fachbereich: **frz (1954)**

Angemeldet als **Ralph Hafner** » [Abmelden](#) | [Hilfe](#)

Neue Einträge

SIS Schlagwort	Formal- stelle	Zeit- schlag- wort	Sach- begriff	Geogra- fikum	Werk	Veran- staltung	Organi- sation	Person (individua- liert)	Familie	Quelle	
Apokryphen	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/> GND <input type="checkbox"/> GND						
A quatre pas du soleil	<input type="checkbox"/>	<input type="checkbox"/>									
arriver	<input type="checkbox"/>	<input type="checkbox"/>				<input type="checkbox"/> GND					
Artamene	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/> GND						
Art d'amours	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> GND <input type="checkbox"/> GND	<input type="checkbox"/> GND <input type="checkbox"/> GND	<input type="checkbox"/> GND <input type="checkbox"/> GND	<input type="checkbox"/> GND <input type="checkbox"/> GND	<input type="checkbox"/> GND <input type="checkbox"/> GND	<input type="checkbox"/> GND <input type="checkbox"/> GND			
Arthur	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> GND <input type="checkbox"/> GND <input type="checkbox"/> GND <input type="checkbox"/> GND	<input type="checkbox"/> GND <input type="checkbox"/> GND <input type="checkbox"/> GND <input type="checkbox"/> GND	<input type="checkbox"/> GND <input type="checkbox"/> GND <input type="checkbox"/> GND	<input type="checkbox"/> GND <input type="checkbox"/> GND <input type="checkbox"/> GND	<input type="checkbox"/> GND <input type="checkbox"/> GND <input type="checkbox"/> GND	<input type="checkbox"/> GND <input type="checkbox"/> GND <input type="checkbox"/> GND <input type="checkbox"/> GND			
				<input type="checkbox"/> GND	<input type="checkbox"/> GND		<input type="checkbox"/> GND	<input type="checkbox"/> GND			

1 - 10 von 10 Einträgen

Bearbeitete Einträge

SIS Schlagwort	Formal- stelle	Zeit- schlag- wort	Sach- begriff	Geogra- fikum	Werk	Veran- staltung	Organi- sation	Person (individua- liert)	Familie	Quelle	
----------------	-------------------	--------------------------	------------------	------------------	------	--------------------	-------------------	---------------------------------	---------	--------	--

SiGMaMat – Screenshot

SIS-Visualizer

Systematikbrowser

Visualisiert

- die Hierarchien und
- Strukturfehler



Viewer

Suchen

Fachgebiete:

sci
ska - sks
sla - slx
soz
spa
spk
spp
spr
ssl
sta
stq
swa

- ▼ sla - slx = Slawistik
 - ▶ sla = Slawische Sprachen / Philologie
 - ▶ slb = Altkirchenslawisch / Philologie
 - ▶ slc - slk = Südslawische Sprachen / Philologie
 - ▼ sll - slo = Ostslawische Sprachen / Philologie
 - ▶ sll = Ostslawische Sprachen / Allgemein / Philologie
 - ▼ slm = Russisch / Philologie
 - ▶ slm 3 - slm 299 = Russisch / Linguistik
 - ▶ slm 300 - slm 309 = Russisch / Philologie
 - ▼ slm 310 - slm 999 = Russisch / Literaturwissenschaft
 - ▶ slm 310 - slm 679 = Russisch / Literatur / Theorie
 - ▼ slm 680 - slm 999 = Russisch / Literatur / Historischer Teil
 - ▶ slm 680 - slm 699 = Russisch / Literatur / Geschichte / Gesamtzeitraum
 - ▶ slm 700 - slm 819 = Russisch / Literatur / Mittelalter
 - ▼ slm 820 - slm 999 = Russisch / Literatur / Neuzeit
 - ▶ slm 820 - slm 839 = Russisch / Literatur / Neuzeit
 - ▶ slm 840 - slm 859 = Russisch / Literatur / Geschichte 1400-1500
 - ▶ slm 860 - slm 879 = Russisch / Literatur / Geschichte 1500-1600
 - ▶ slm 880 - slm 899 = Russisch / Literatur / Geschichte 1600-1700
 - ▶ slm 900 - slm 919 = Russisch / Literatur / Geschichte 1700-1800
 - ▶ slm 919.10 = Russisch / Literarische Zeitschrift / Geschichte 1700-1800
 - ▶ slm 920 - slm 939 = Russisch / Literatur / Geschichte 1800-1900
 - ▶ slm 939.10 = Russisch / Literarische Zeitschrift / Geschichte 1800-1900
 - ▶ slm 940 - slm 959.99 = Russisch / Literatur / Geschichte 1900-2000
 - ▶ slm 960 - slm 979 = Russisch / Literatur / Geschichte 2000-2100
 - ▶ sln = Ukrainisch / Philologie
 - ▶ slo = Weißrussisch / Philologie
 - ▶ slp - slx = Westslawische Sprachen / Philologie

Systemstellenfilter:

s

Go

Informationen

Systemstelle: slm 940 - slm 959.99

Standort: G 3a

Erfassungsdatum: 17.07.2000

Korrekturdatum: 19.08.2014

[In SIS anzeigen](#)

Schlagwortkette

Russisch / Literatur /
Geschichte 1900-2000

Paralleleinträge

Russisch / Literaturgeschichte
1900-2000

SIS-Visualizer – Fachbeispiel: Slavistik

Nächste Schritte



<http://upload.wikimedia.org/wikipedia/commons/f/f6/HAL9000.svg>

KoKon: Beispiele

Mögliche semantische Verschränkungen von Klassifikationen

– Trivialstbeispiele

KonSys: Sprachwissenschaft Morphologie Wortbildung (spr 165 - spr 173)

RVK: Sprachwissenschaft Morphologie Wortbildung (ET 350)

KonSys: Sprachwissenschaft: Syntax / Semantik (spr 181)

RVK: Sprachwissenschaft: Syntax und Semantik (ET 750)

– Trivialbeispiel

KonSys: Rechtswissenschaft / Linguistik (spr 28.85)

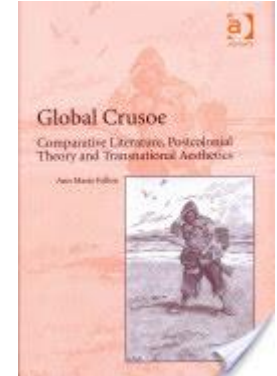
RVK: Forensische Linguistik (ES 970)

– komplexes Beispiel

KonSys: Literaturwissenschaft / Poetik / Gattungsgeschichte /
Prosa / Epik / Roman / Abenteuerroman (lit 464)

DDC: Geschichte, Darstellung, kritische Bewertung von mehr als
zwei Literaturen / Abenteuerprosa—Geschichte und
Literaturwissenschaft (DDC 809.387)

Auto-SE: Beispiel (1/3)



**Fallon, Ann Marie: Global Crusoe.
Comparative literature, postcolonial
theory and transnational aesthetics.**

Farnham [u.a.]: Ashgate 2011.

ISBN: 978-1-4094-2998-2



SE-Daten sammeln

Auto-SE: Beispiel (2/3)

Verschränken – Finden der Beziehungen zwischen den SE-Systemen

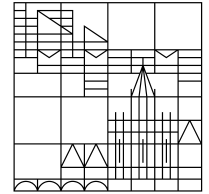
- GND-SW-WT: Defoe, Daniel / Robinson Crusoe GND ID: 4281761-4 →
 1. eng 919:d314:kr61 = Defoe, Daniel / Robinson Crusoe / Ausgabe ODER
 2. eng 919:d314:yk61 = Defoe, Daniel / Robinson Crusoe / Sekundärliteratur
 3. lit 770:r65 = Robinson CrusoeEntscheidung anhand von Kontext

Auto-SE: Beispiel (3/3)

- DDC: 809.387 →
GND Abenteuerroman 4000089-8 →
lit 464 = Abenteuerroman +
deu 588:a14 = Deutsch / Abenteuerroman / Geschichte
eng 588:a14 = Englisch / Abenteuerroman / Geschichte
frz 588:a14 = Französisch / Abenteuerroman / Geschichte
spa 588:a14 = Spanisch / Abenteuerroman / Geschichte

Literatur

- Bösing, Laurenz; Stoltzenburg, Joachim; Thomashoff, Barbara (1969): Regeln für den Aufbau von Buchsignaturen. (Überarbeitete und auf den neuesten Stand gebrachte Fassung der Regeln vom April 1967 (B/Sto/Th) unter Berücksichtigung aller späteren Anhänge und Ergänzungen). Konstanz: Bibliothek der Universität Konstanz (Bibliothek aktuell / Sonderheft, 1).
- Kasprzik, Anna (2013): Projektbericht: Implementierung eines Hierarchisierungsalgorithmus' für die Konstanzer Systematik. Bibliothek der Universität Konstanz. Online verfügbar unter <http://nbn-resolving.de/urn:nbn:de:bsz:352-241667>.
- Lorenz, Bernd (2003): Systematische Aufstellung in Vergangenheit und Gegenwart. Wiesbaden: Harrassowitz (Beiträge zum Buch- und Bibliothekswesen, 45).
- Müller-Dreier, Armin (1994): Einheitsklassifikation. Die Geschichte einer fortwirkenden Idee. Wiesbaden: Harrassowitz (Beiträge zum Buch- und Bibliothekswesen, 35).
- Schelling, Bernd (2014): KoKon – Kontextsensitiver Abgleich für Klassifikationen, unpublished



Herzlichen Dank!

KIM – IT- und Bibliotheksdienste

Ralph Hafner

Fachreferent, Sacherschließungsreferent

ralph.hafner@uni-konstanz.de

Bernd Schelling

Abteilungsleiter Contentbasierte Dienste

bernd.schelling@uni-konstanz.de