

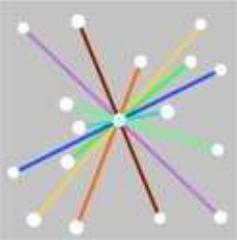
Vom Stringmatching zur Begriffsexploration: das Potential integrierter begrifflicher Interoperabilität

Jessica Hubrich
DFG-Projekt CrissCross

Fachhochschule Köln
Institut für Informationsmanagement

12. Tagung der Deutschen ISKO – *“Wissen – Wissenschaft - Organisation”*.
20. Oktober 2009

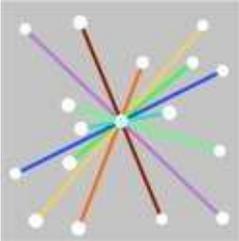
Übersicht



Vom Stringmatching zur Begriffs-
exploration: semantische Interoperabilität

- Ausgangssituation: Thematische Suche
- Stufenmodell semantischer Interoperabilität
 - Wortbasierte Interoperabilität
 - Begriffliche Interoperabilität
 - Differenzierte Interoperabilität
- Datenhaltungsmodelle
 - Externe Datenhaltung
 - Integrierte Datenhaltung
- Exploration mittels integrierter begrifflicher Interoperabilität
 - Erweiterte Begriffsexploration
 - Umfassende Begriffsexploration

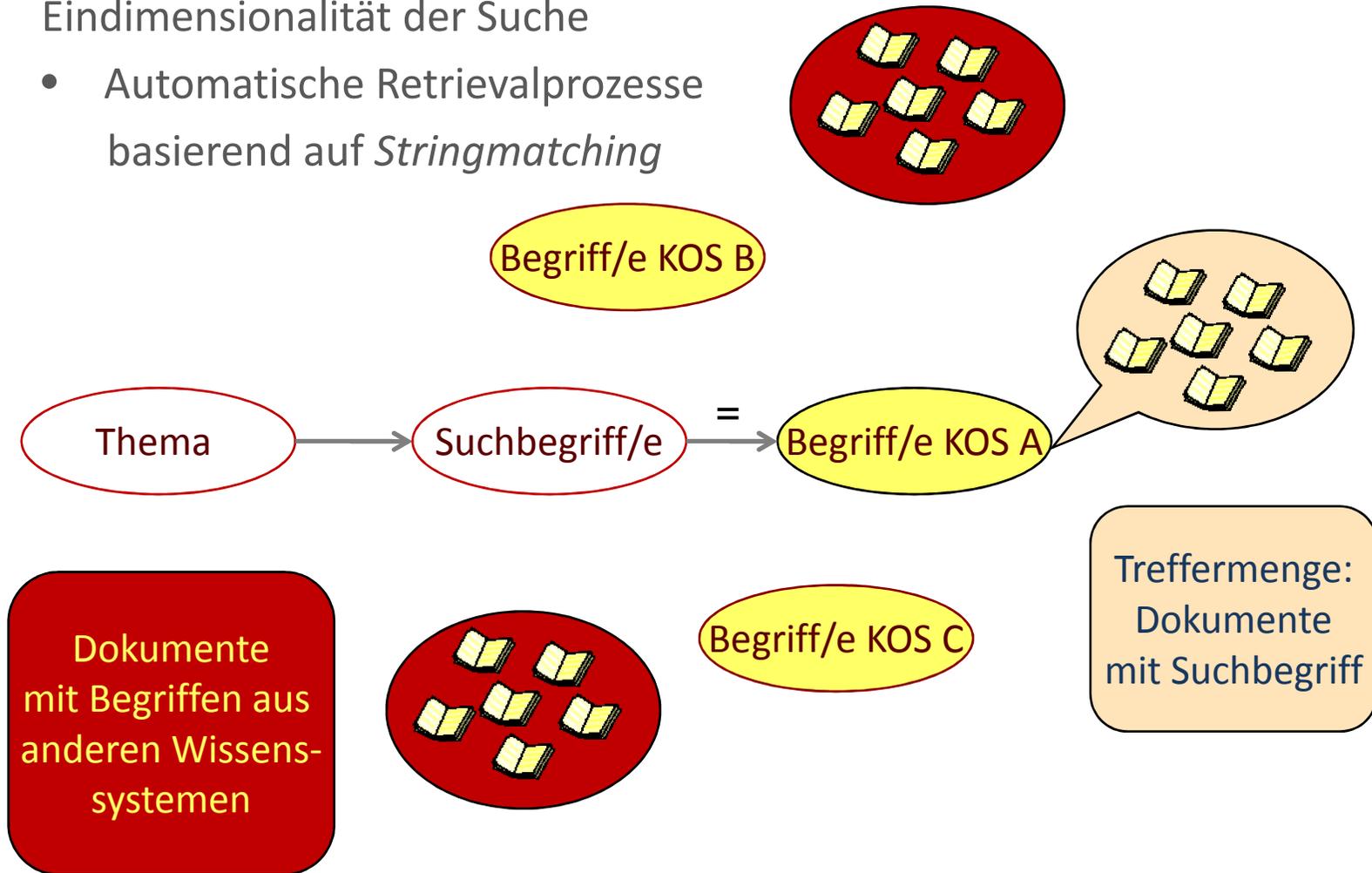
Thematische Suche – Grundvorstellung



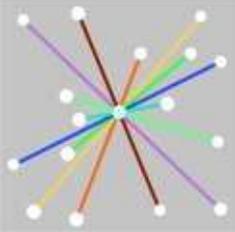
Vom Stringmatching zur Begriffs-
exploration: semantische Interoperabilität

Eindimensionalität der Suche

- Automatische Retrievalprozesse basierend auf *Stringmatching*



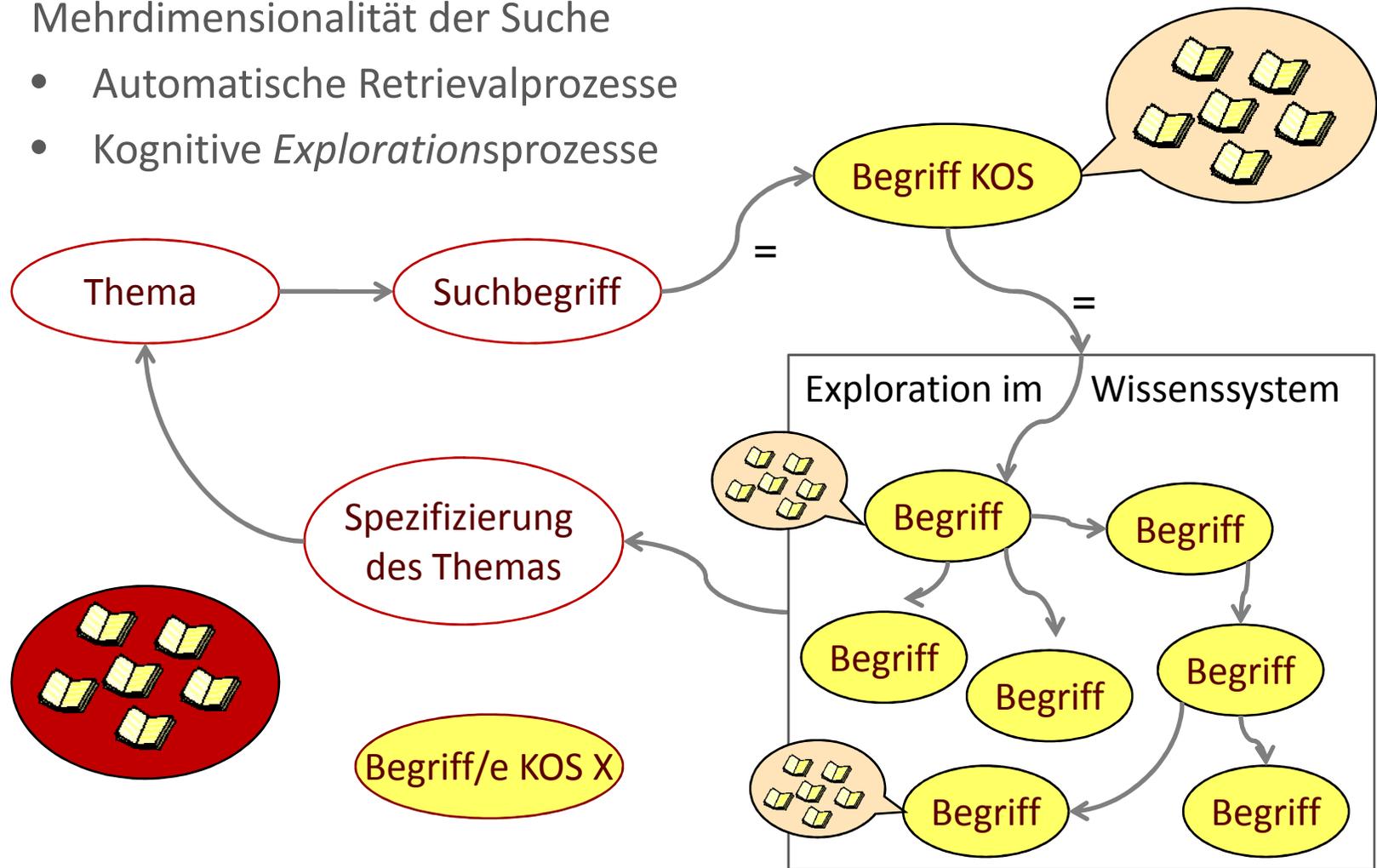
Thematische Suche — Erweitertes Verständnis



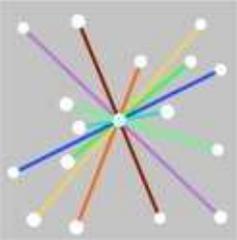
Mehrdimensionalität der Suche

- Automatische Retrievalprozesse
- Kognitive *Explorations*prozesse

Vom Stringmatching zur Begriffs-
exploration: semantische Interoperabilität



Modelle semantischer Interoperabilität



Vom Stringmatching zur Begriffs-
exploration: semantische Interoperabilität

Stufenmodell semantischer Interoperabilität

Grundlagen

- Mappingstrategien
- Eigenschaften zur Unterstützung thematischer Suchprozesse

3 – Differenzierte Interoperabilität

2 – Begriffliche Interoperabilität

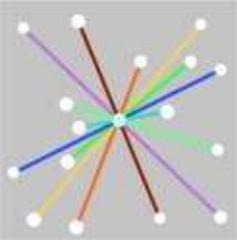
1 – Wortbasierte Interoperabilität

Datenhaltungsmodelle

Externe Datenhaltung

Integrierte Datenhaltung

Interoperabilitätsstufen 1 - 2



Vom Stringmatching zur Begriffs-
exploration: begriffliche Interoperabilität

1 – Wortbasierte Interoperabilität

- Einsatz automatischer Verfahren der Heterogenitätsbehandlung
- Wörter/Zeichenketten als Grundlage
- Primär One-to-One Mapping

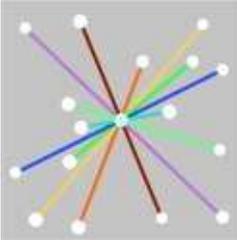
2 – Begriffliche Interoperabilität

- Intellektuelle Methoden der Heterogenitätsbehandlung
- Bedeutungsumfang von Begriffen als Grundlage
- Grundvoraussetzung für präzise begriffliche Suchen
- Evtl. Direktionalität; One-to-One + One-to-Many Mapping

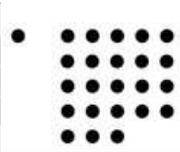
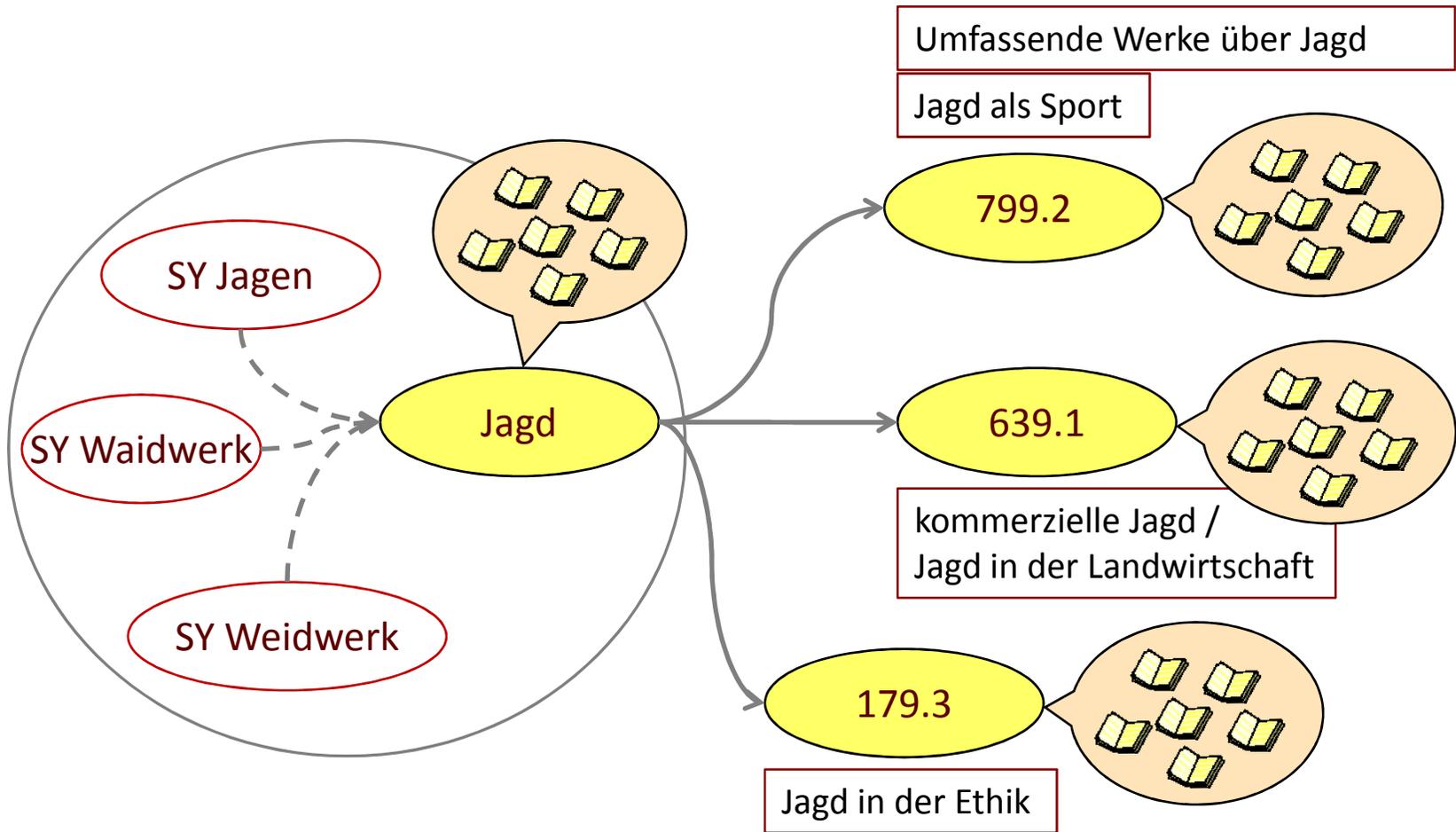
→ Erweiterung der sprachlichen Basis für Stringmatching-Ansätze

→ Erhöhung von Recall und Precision

Interoperabilitätsstufe 2 – CrissCross

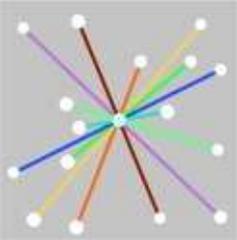


Vom Stringmatching zur Begriffs-
exploration: begriffliche Interoperabilität



Fachhochschule Köln
Cologne University of Applied Sciences

Interoperabilitätsstufe 3

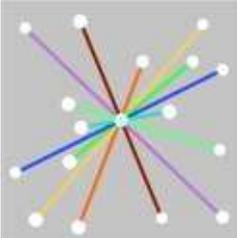


Vom Stringmatching zur Begriffs-
exploration: begriffliche Interoperabilität

3 – Differenzierte Interoperabilität

- Spezifizierung der Art der Beziehung zwischen Begriffen unter Berücksichtigung sprachlicher, struktureller und typologischer Unterschiede
 - Ambitionierte Ansätze erfordern ein über die üblichen Basisrelationen (OB, UB, VB) hinaus gehendes Inventar
 - Ausweisung von logischen Eigenschaften zum Ziehen automatischer Inferenzen
 - Typisierung von Relationen
 - Idealerweise Entwicklung eines allgemeinen Inventars an Mappingrelationen zugunsten allgemeiner Anwendbarkeit
- Grundlage für effiziente Unterstützung von Prozessen der Begriffs- und Wissensexploration in heterogenen Informationsräumen
- Grundlage für Rankingmechanismen

Interoperabilitätsstufe 3 – CrissCross – Ranking

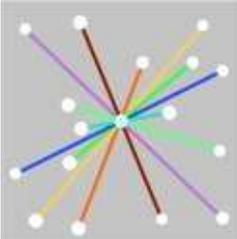


4 *Determiniertheitsgrade*, die die Beziehung zwischen SWD-Schlagwort und DDC-Klasse abbilden und sich an Thema-Klasse-Beziehungen der DDC orientieren, bieten Möglichkeiten zum Ranking von Begriffen bzw. Dokumenten

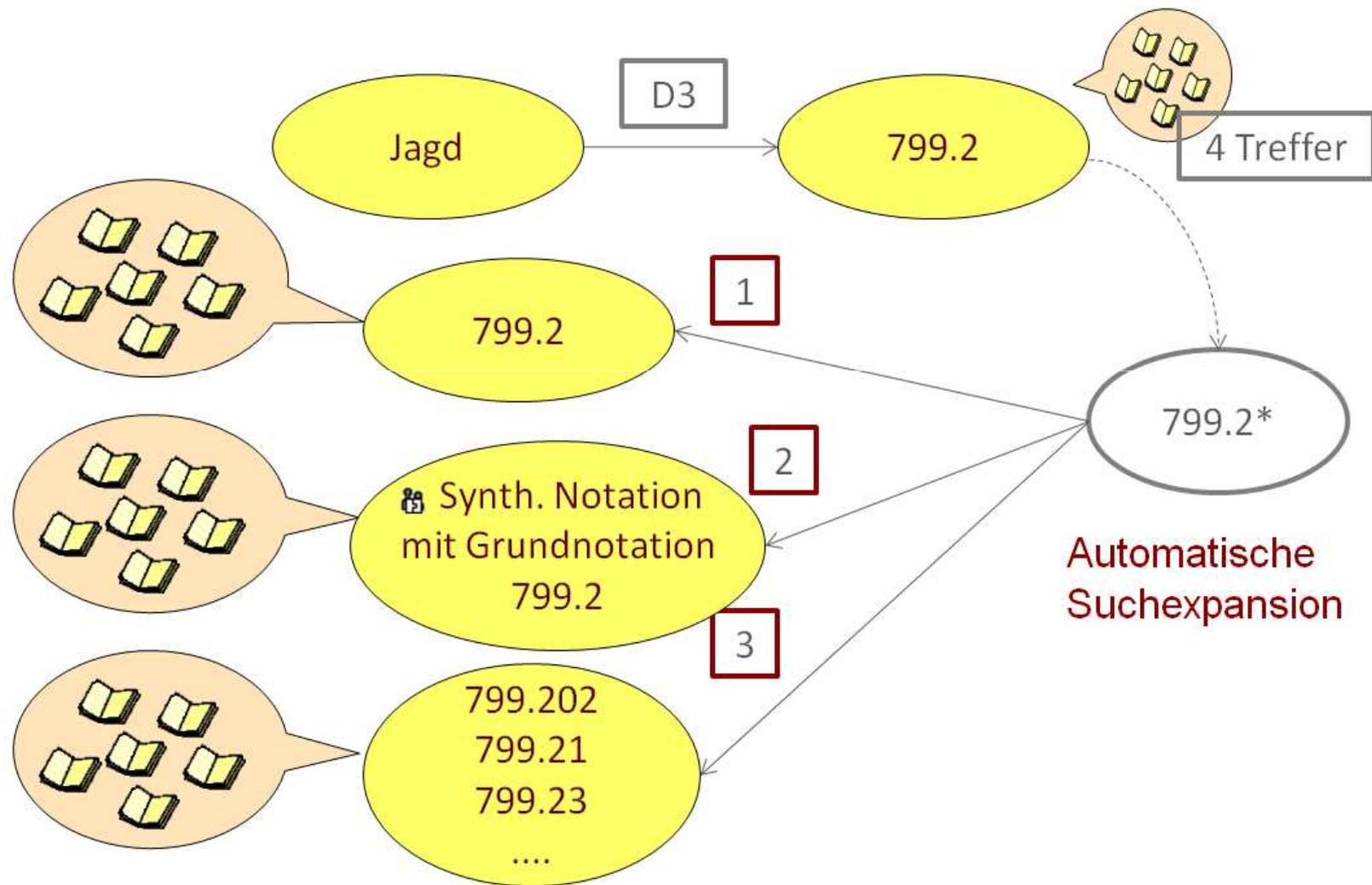
Vom Stringmatching zur Begriffs-
exploration: semantische Interoperabilität

		SWD-Schlagwörter	Relevanz
<u>790</u>	<u>Sport, Spiele, Unterhaltung</u>		
<u>799</u>	<u>Fischfang, Jagd, Schießen</u>	<u> s Fallenjagd</u>	hoch
799.2	Jagd	<u> s Fangjagd</u>	
<u>799.202</u>	<u>Verschiedenes</u>	<u> s Jagd</u>	
<u>799.206</u>	<u>Organisationen und Management</u>		
<u>799.2074</u>	<u>Jagdmuseen</u>	<u> s Abschuss</u>	mittel
<u>799.2076</u>	<u>Prüfen und Üben</u>	<u> s Abschussplan</u>	
<u>[799.209]</u>	<u>Historische, geografische, personenbezogene Behandlung</u>	<u> s Ansitz</u>	
		<u> s Ansprechen</u>	
<u>799.21-799.23</u>	<u>Methoden</u>	<u> s Balg</u>	
<u>799.24-799.27</u>	<u>Jagd auf einzelne Arten von Tieren</u>	<u> s Beutefang</u>	
<u>799.29</u>	<u>Historische, geografische, personenbezogene Behandlung</u>	<u> s Blatten <Jagd></u>	
		<u> s Deckung <Schutz></u>	
		<u> s Gesellschaftsjagd</u>	
		<u> s Jagdkleidung</u>	

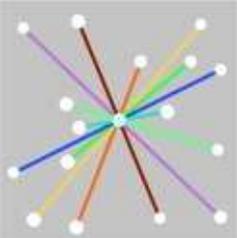
Interoperabilitätsstufe 3 – CrissCross – Ranking



Vom Stringmatching zur Begriffs-
exploration: semantische Interoperabilität



Datenhaltungsmodelle



Vom Stringmatching zur Begriffs-
exploration: semantische Interoperabilität

Datentechnische Rahmenbedingung für dauerhafte Tauglichkeit:
Verbindung nicht über Klartextbezeichnungen, sondern auf der Grundlage
von eindeutigen Identifikationsmerkmalen (IDs, URIs)

Externe Datenhaltung

Datenhaltung
KOS 1

Datenhaltung
Mapping KOS 1 - KOS 2

Datenhaltung
KOS 2

Eingabe: 1250:01-07-88 Änderung: 9999:25-04-09 19:50:51 Status

Integrierte Datenhaltung

005 Ts1

011 /s

012 /w

021 4027955-8

800 **|s|Jagd**

808 |a|M

808 |d||Keine weitere Kompositabildung mit einzelnen Tiernamen

809 |x|be;me *erl

810 32.9

816 **179.3#1# [2007-01-01]**

816 **346.0469549#1# [2007-01-01]**

816 **639.1#3# [2007-01-01]**

816 **799.2#3# [2007-01-01]**

830 |s|Jagen

830 |s|Waidwerk

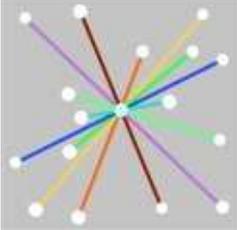
830 |s|Weidwerk

860 **|s|Tierfang**

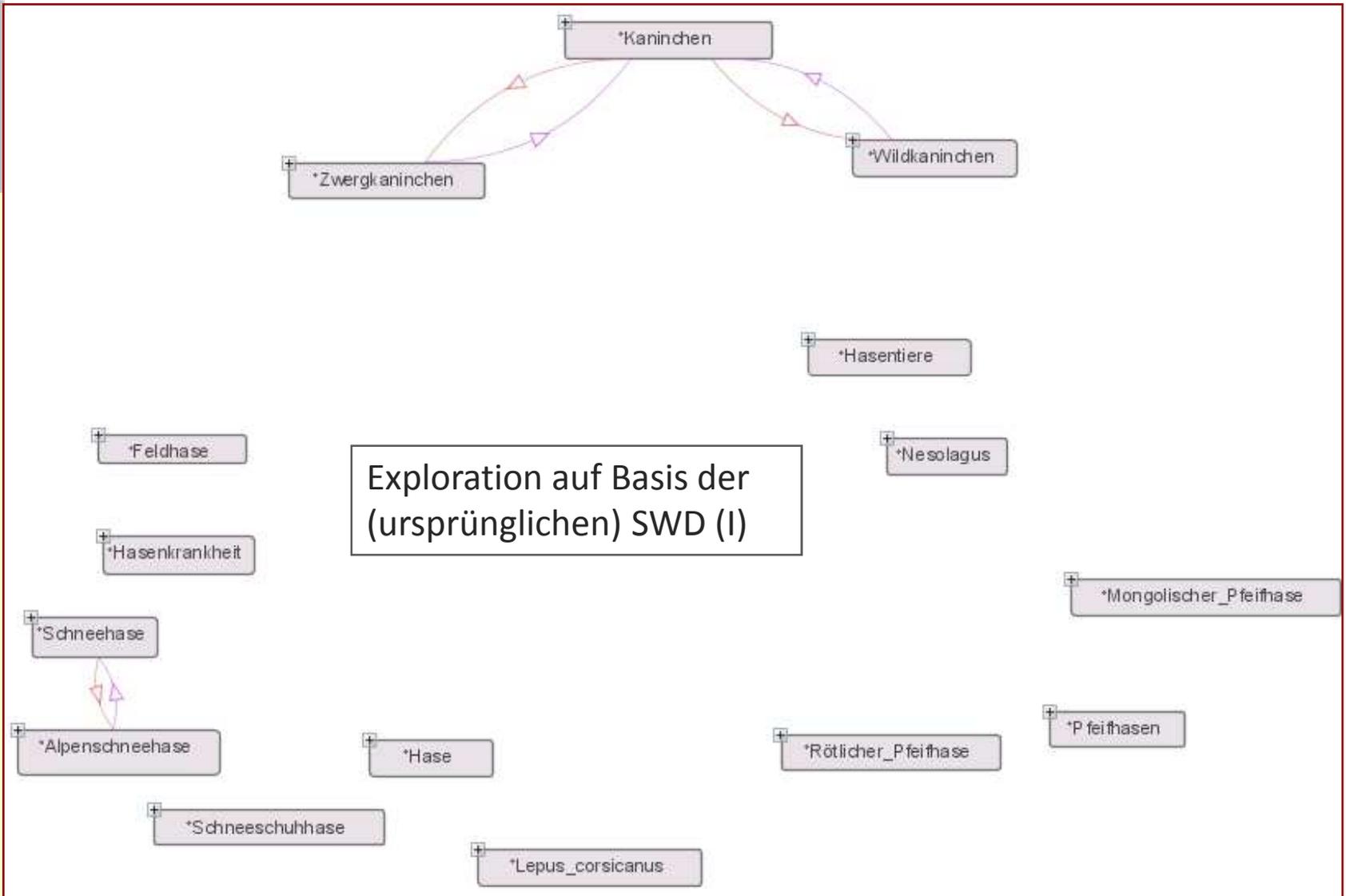
860 **|s|Jagdkunde**

860 **|s|Jäger**

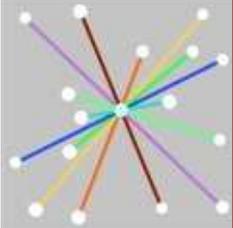
Erweiterte Begriffsexploration (I)



Vom Stringmatching zur Begriffs-
exploration: semantische Interoperabilität

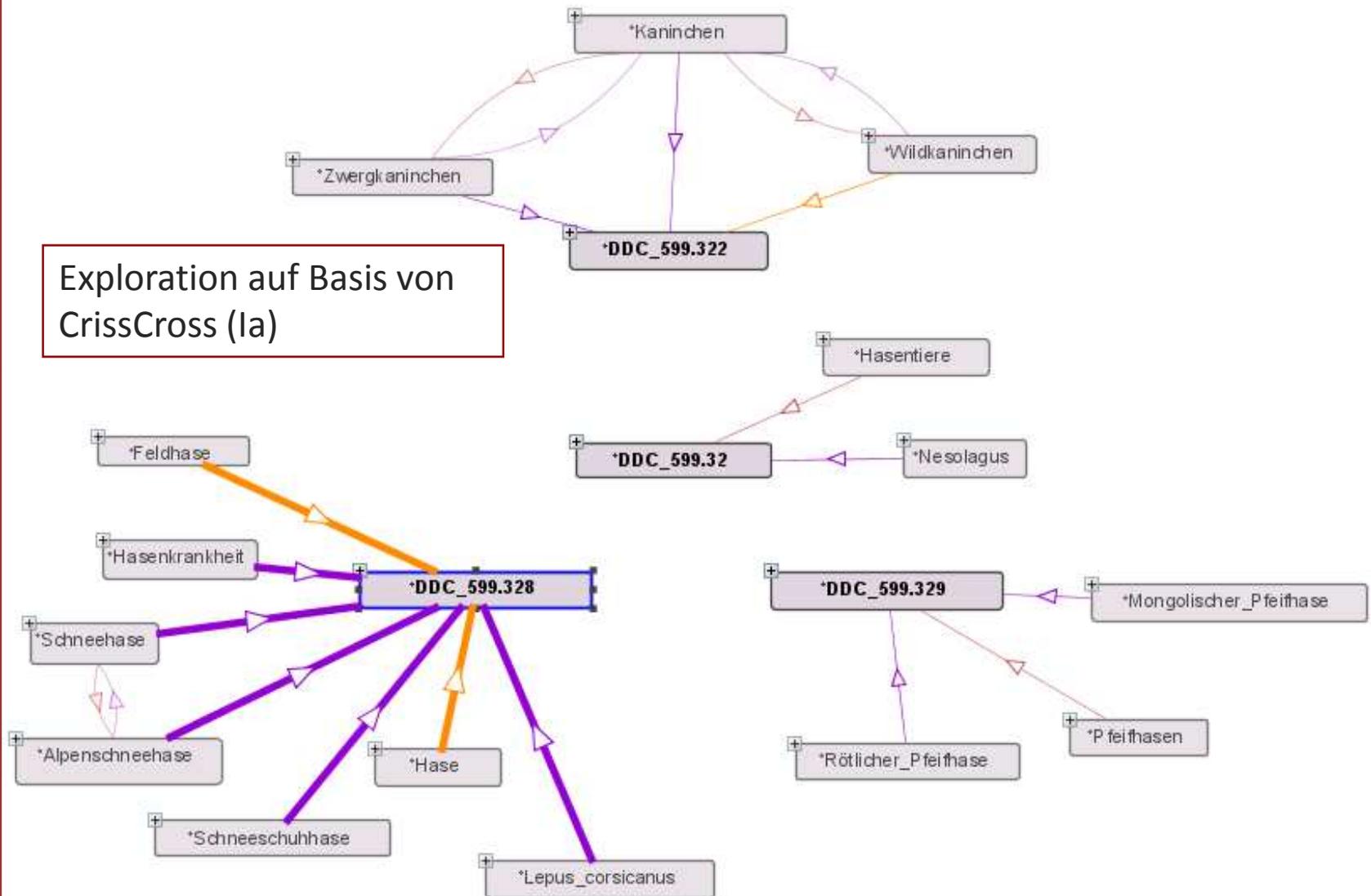


Erweiterte Begriffsexploration (I)

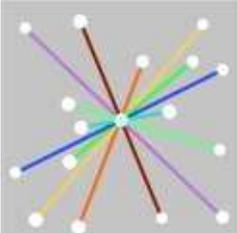


Vom Stringmatching zur Begriffs-
exploration: semantische Interoperabilität

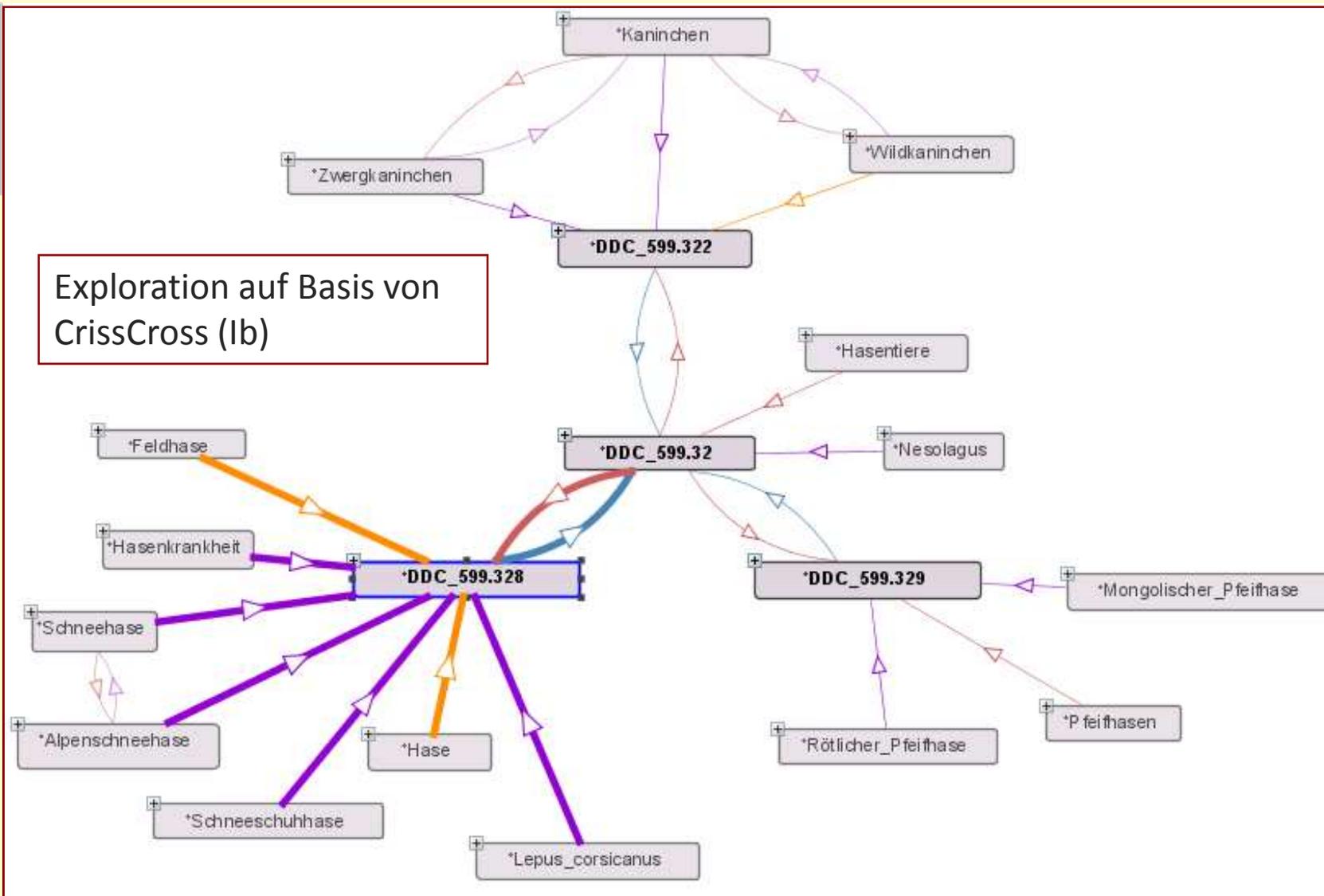
Exploration auf Basis von
CrissCross (Ia)



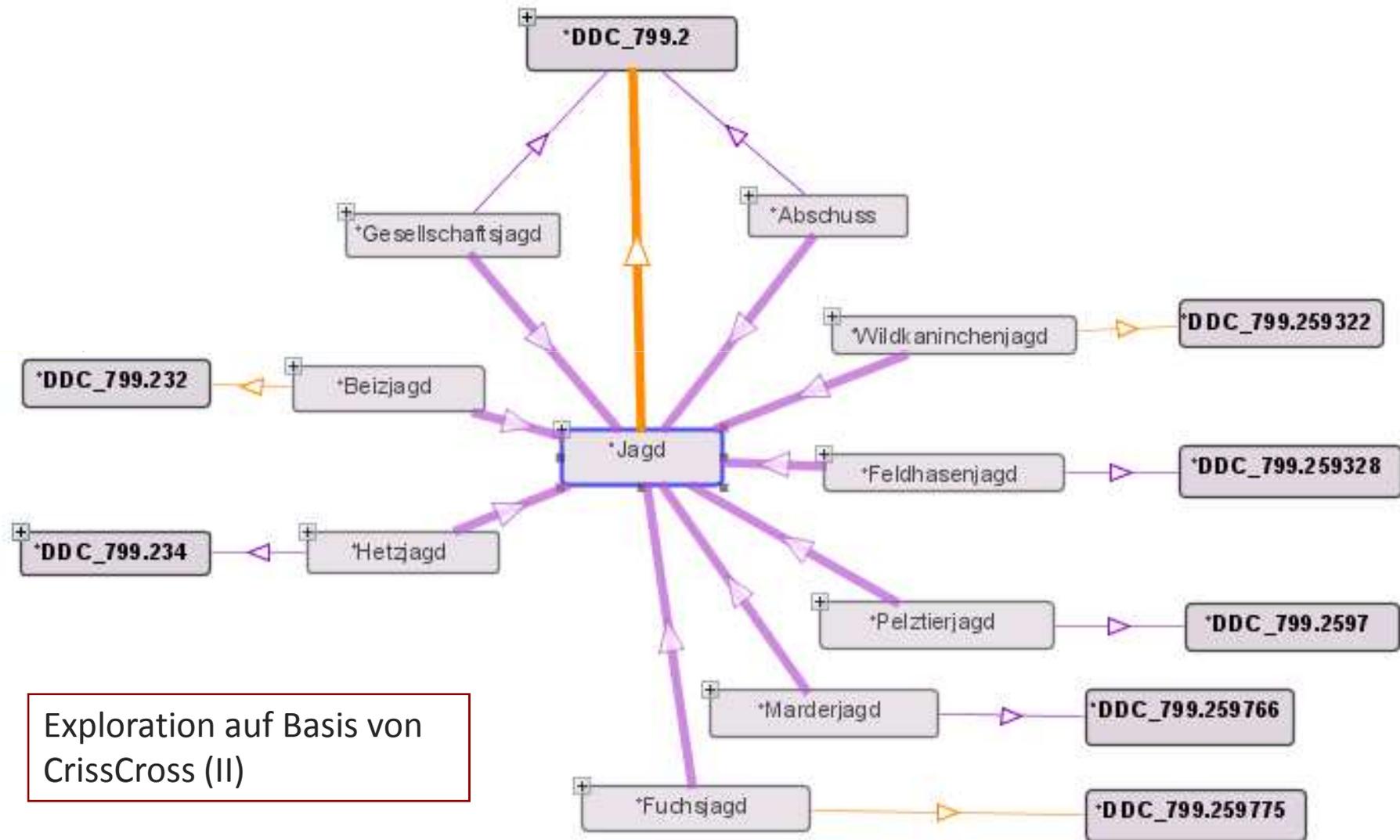
Erweiterte Begriffsexploration (I)



Vom Stringmatching zur Begriffs-
exploration: semantische Interoperabilität

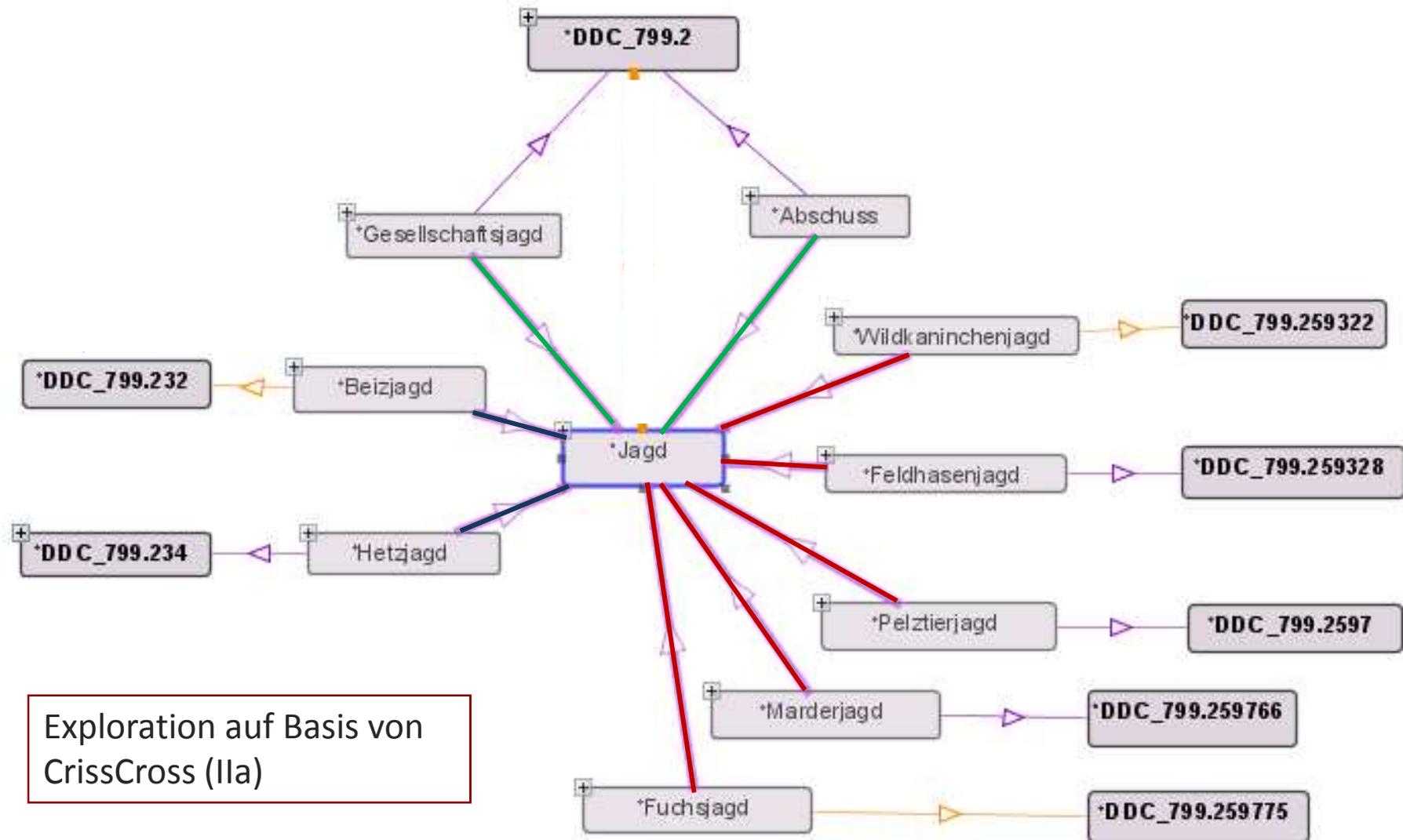


Erweiterte Begriffsexploration (II)

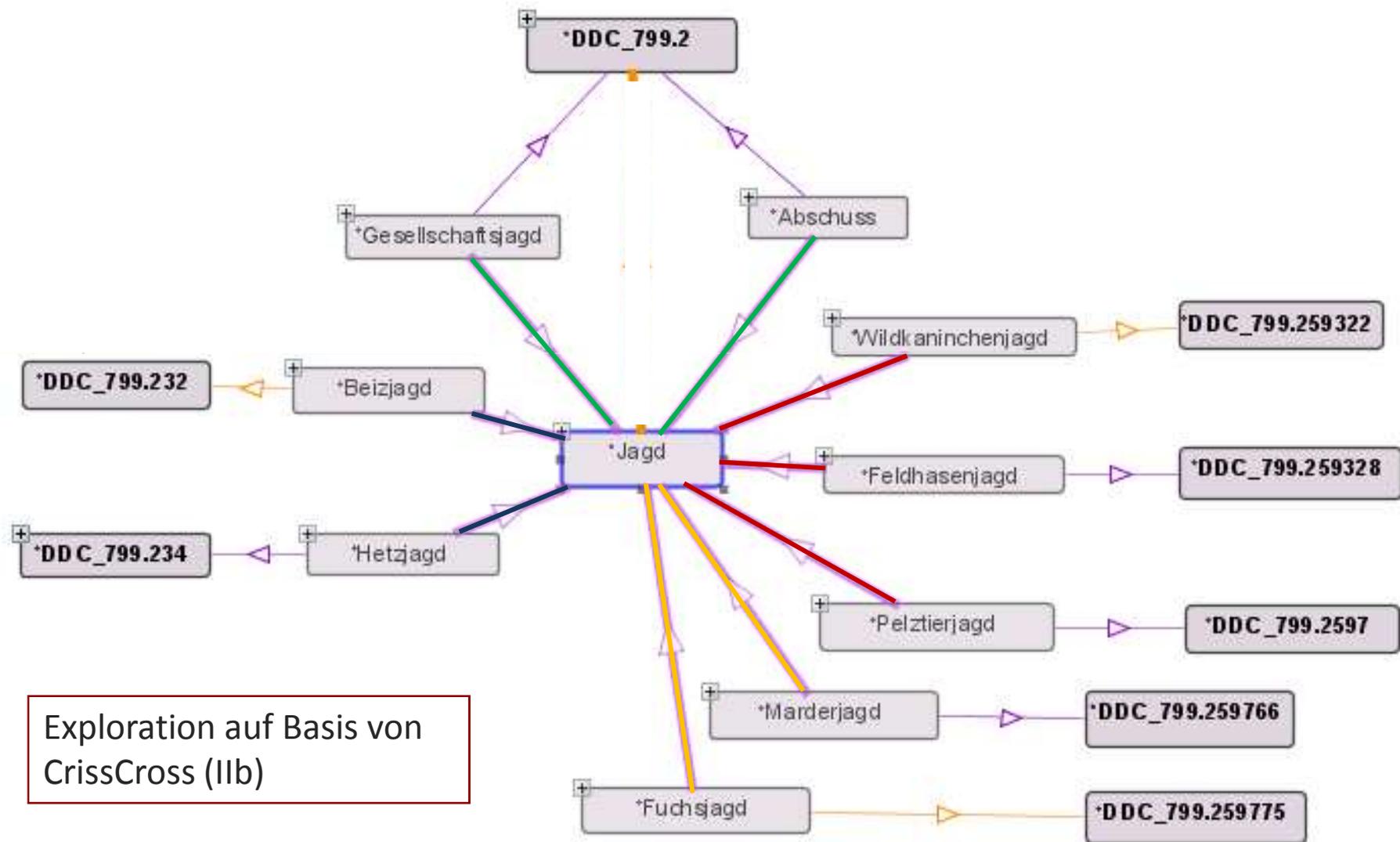


Exploration auf Basis von
CrissCross (II)

Erweiterte Begriffsexploration (II)

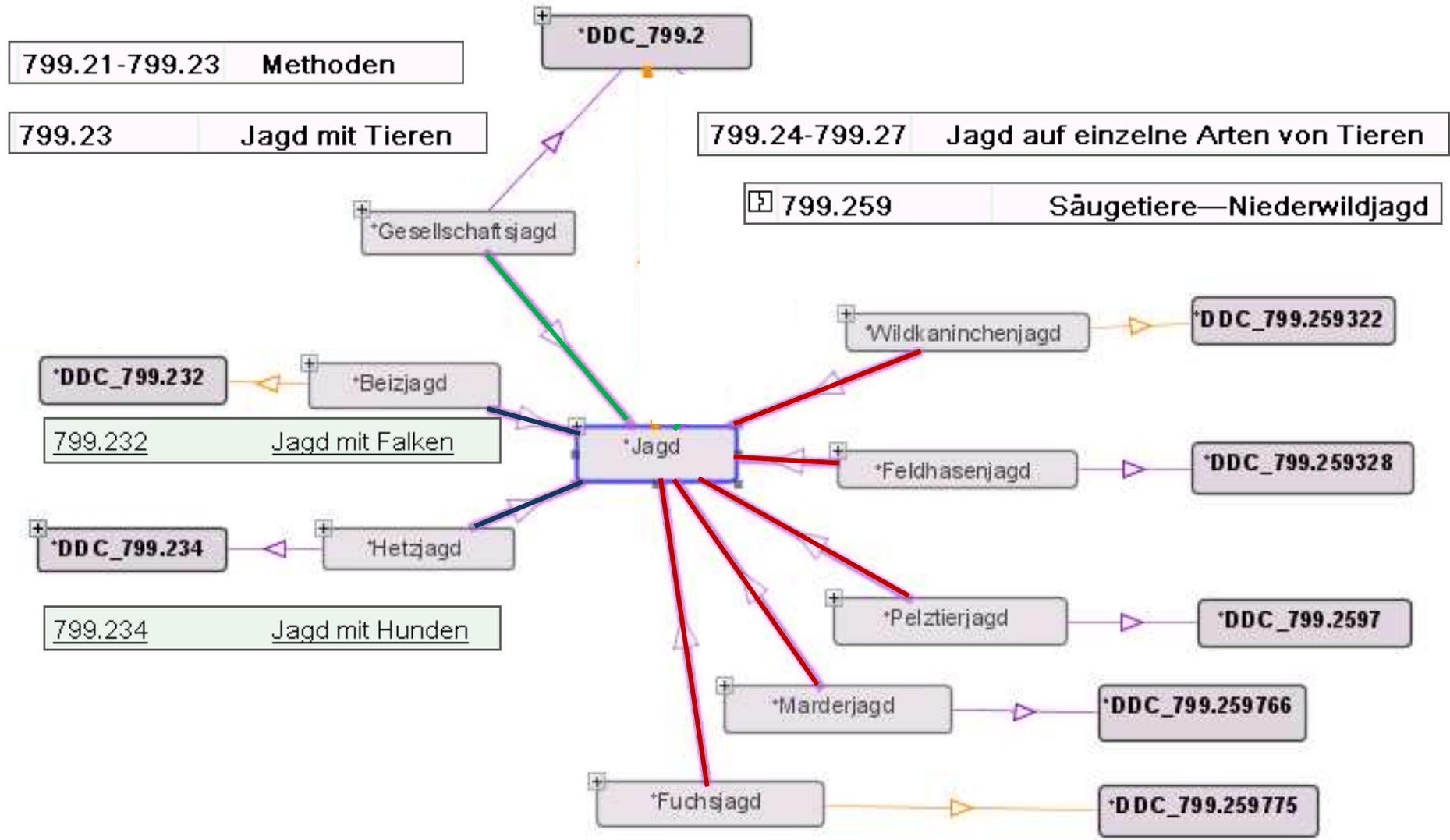


Erweiterte Begriffsexploration (II)

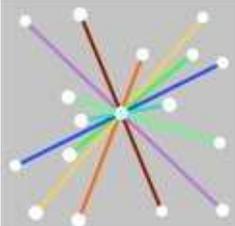


Exploration auf Basis von
CrissCross (IIb)

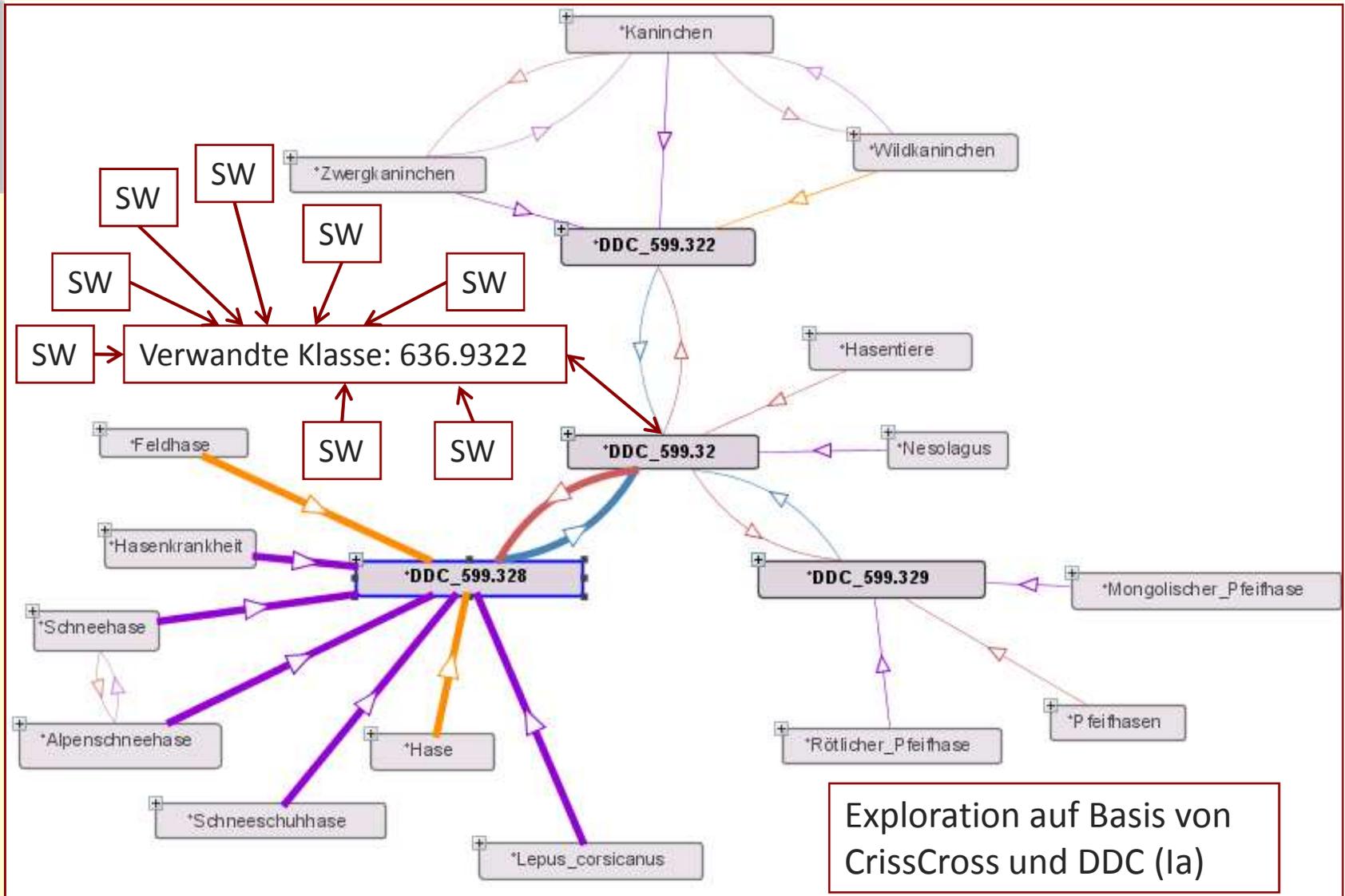
Umfassende Begriffsexploration



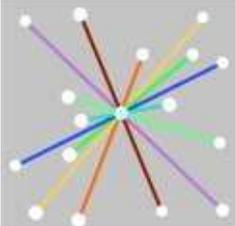
Umfassende Begriffsexploration



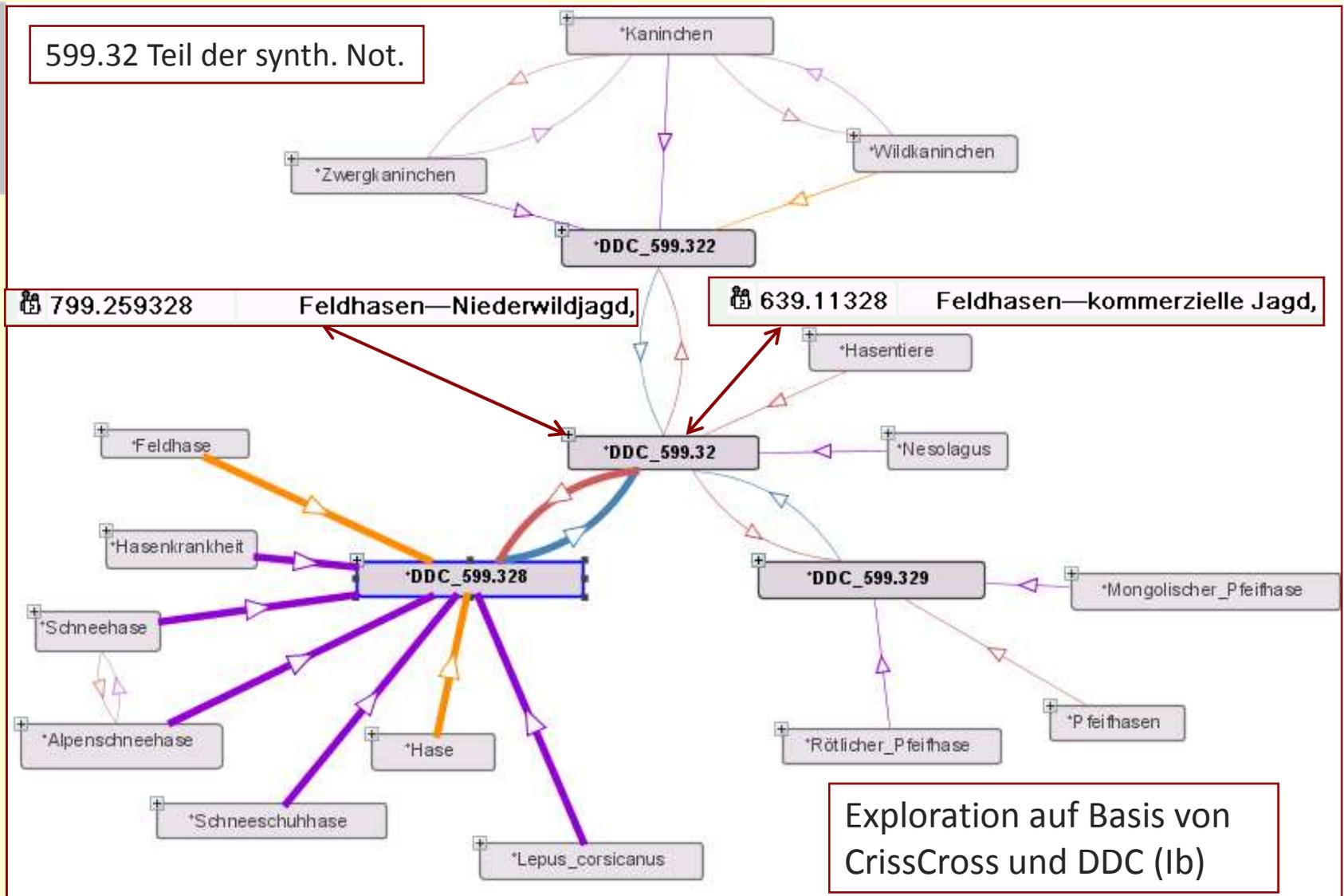
Vom Stringmatching zur Begriffs-
exploration: semantische Interoperabilität



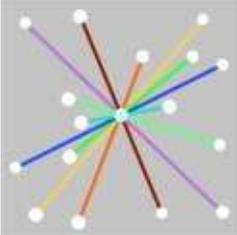
Umfassende Begriffsexploration



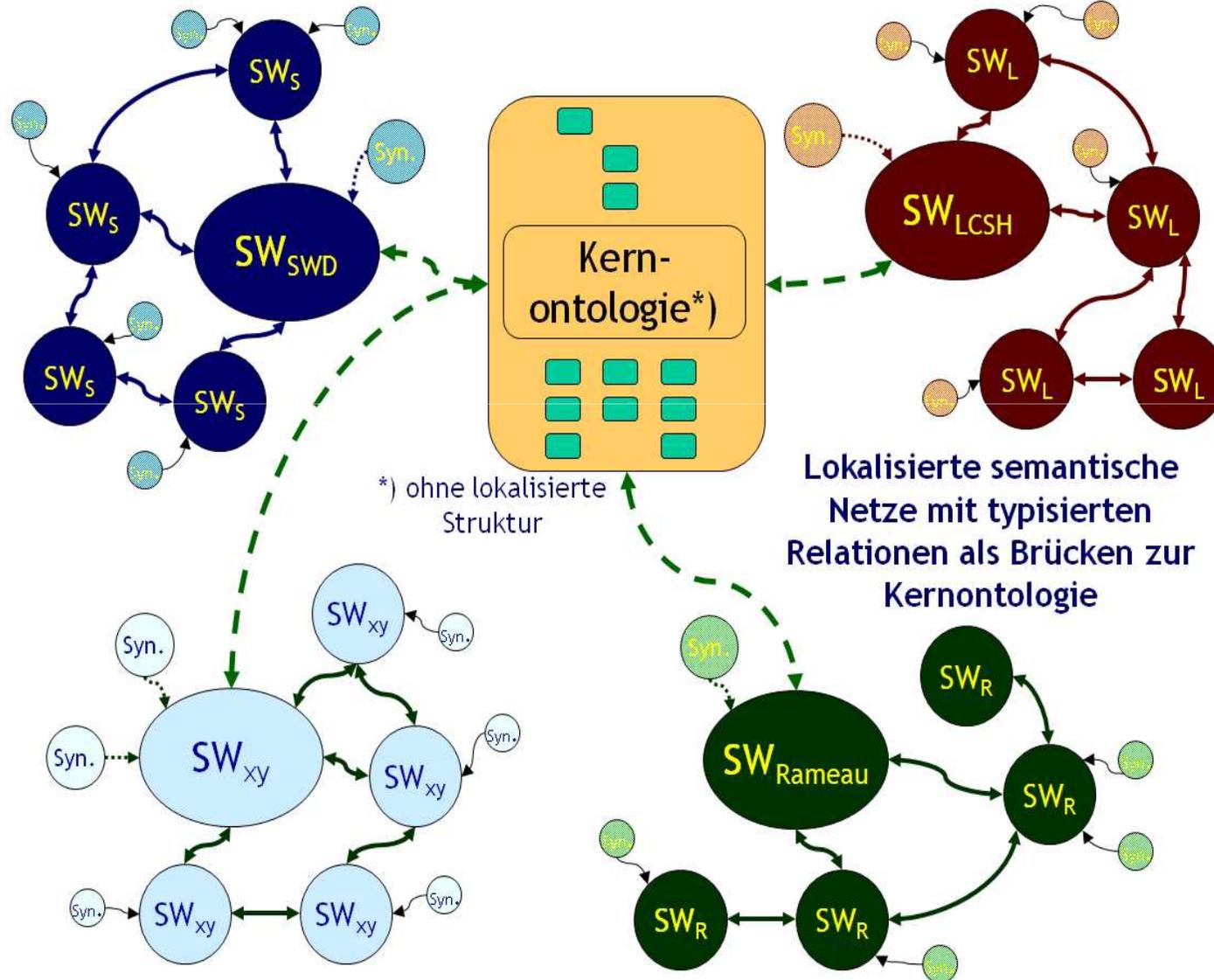
Vom Stringmatching zur Begriffs-
exploration: begriffliche Interoperabilität

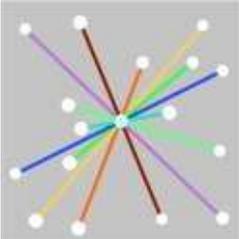


Zukunftsvision?



Vom Stringmatching zur Begriffs-
exploration: semantische Interoperabilität





Vom Stringmatching zur Begriffs-
exploration: semantische Interoperabilität

Vielen Dank für Ihre Aufmerksamkeit!

Webseite DFG-Projekt CrissCross
<http://linux2.fbi.fh-koeln.de/crisscross/index.html>

Projektteam an der Fachhochschule Köln

Prof. Winfried Gödert (Projektleiter)
Jessica Hubrich
Jan-Helge Jacobs
Tina Mengel
Katrin Müller

