

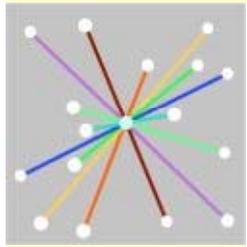
## **Begriffliche Suche und Wissensexploration in heterogenen Informationsräumen**

Jessica Hubrich, M.A., M.L.I.S.  
*Teamleiterin DFG-Projekt CrissCross*

Fachhochschule Köln  
Institut für Informationsmanagement

*30. Österreichischer Bibliothekartag — “The ne(x)t Generation”*  
Inhaltserschließung : Ingredienzen, Instrumentarien, Internationalität  
16. September 2009

# Ausgangssituation / Zielvorstellung



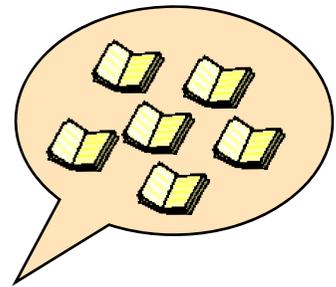
Begriffliche Suche und Exploration  
in heterogenen Informationsräumen

Ausgangssituation

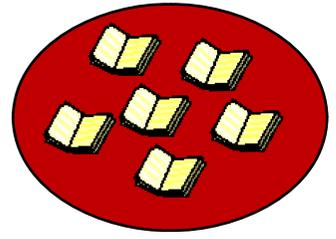
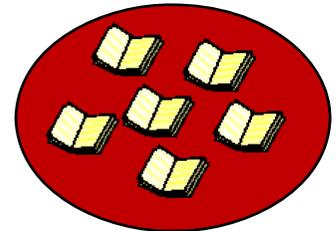
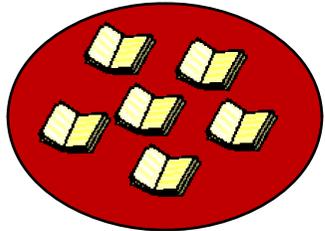
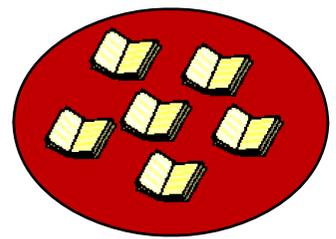
Suchbegriff

=

Begriff KOS A

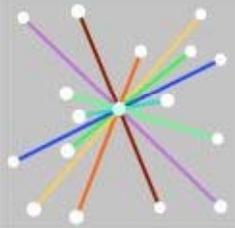


Treffermenge:  
Dokumente  
mit Suchbegriff



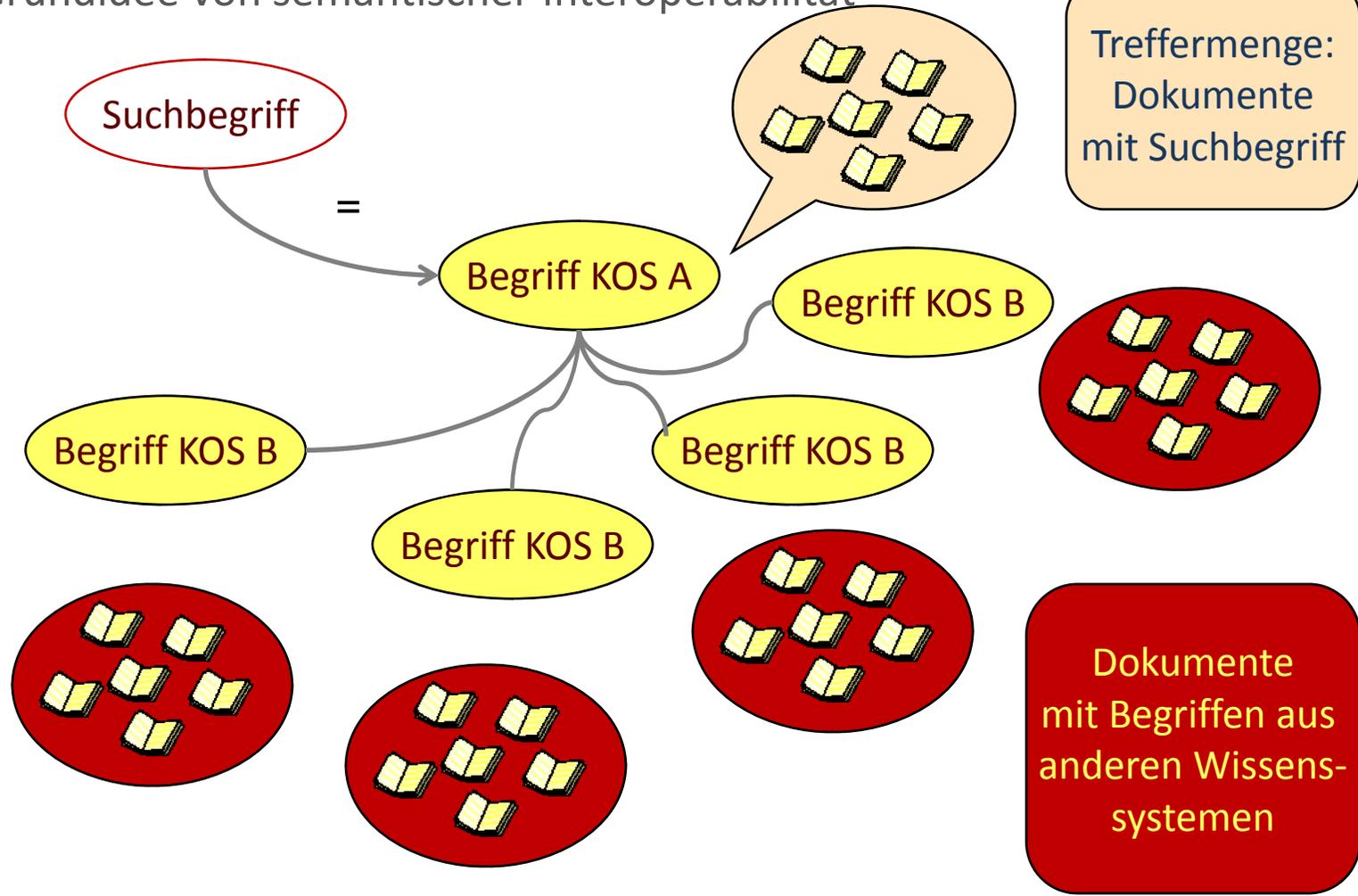
Dokumente  
mit Begriffen aus  
anderen Wissens-  
systemen

# Ausgangssituation / Zielvorstellung

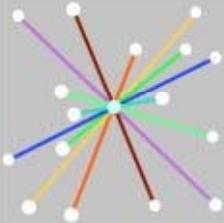


Grundidee von semantischer Interoperabilität

Begriffliche Suche und Exploration  
in heterogenen Informationsräumen



# Ausgangssituation / Zielvorstellung



Zielvorstellung

Suchbegriff

=

Begriff KOS A

Begriff KOS B

Begriff KOS B

Begriff KOS B

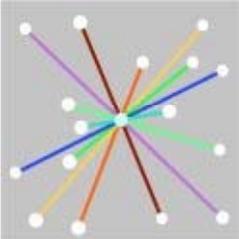
Begriff KOS B

Treffermenge:  
Dokumente  
mit Suchbegriff

Treffermenge:  
Dokumente  
mit Begriffen aus  
anderen Wissens-  
systemen

Begriffliche Suche und Exploration  
in heterogenen Informationsräumen

# Semantische Interoperabilität



Begriffliche Suche und Exploration  
in heterogenen Informationsräumen

Die Idee semantischer Interoperabilität bedarf einer Präzisierung bzgl.

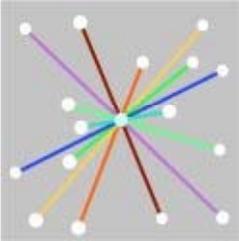
- Art der semantischen Interoperabilität
  - wortbasiert, begrifflich, differenziert
  - Direktionalität
- Art der anvisierten Recherchefunktionalität
  - wortbasiert Suche
  - begriffliche Suche
  - thematische Suche
  - wortbasierte Exploration
  - begriffliche Exploration
  - thematische Exploration

Interdependenz

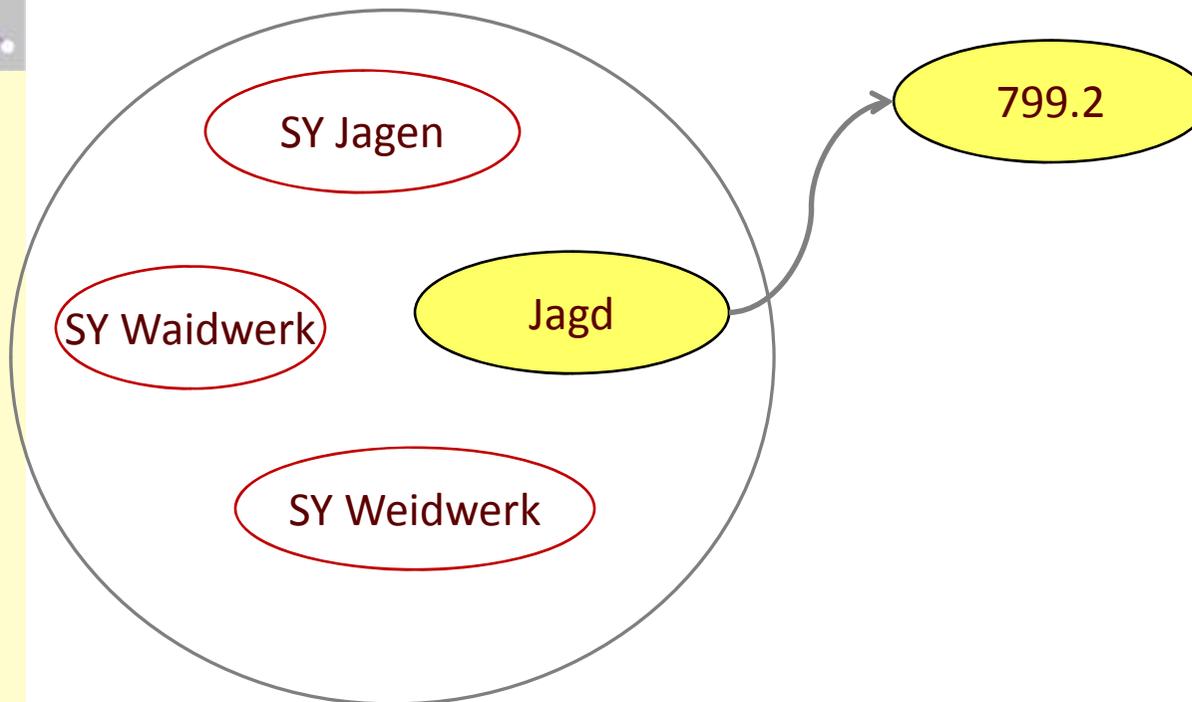
und einer Ergänzung bzgl.

- Strukturierung von Treffermengen / Ranking nach Relevanz
  - *allg. Kriterien:* Ort des Vorkommens (Schlagwortfeld, Titelfeld, Abstract), Häufigkeit des Vorkommens

# CrissCross — Begriffliche Interoperabilität



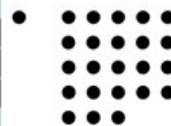
- unidirektional: SWD → DDC



Begriffliche Suche und Exploration  
in heterogenen Informationsräumen

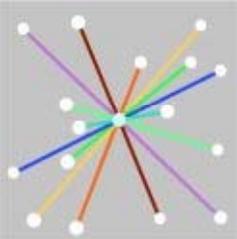
Deutsche  
Forschungsgemeinschaft  
**DFG**

DEUTSCHE  
NATIONAL  
BIBLIOTHEK



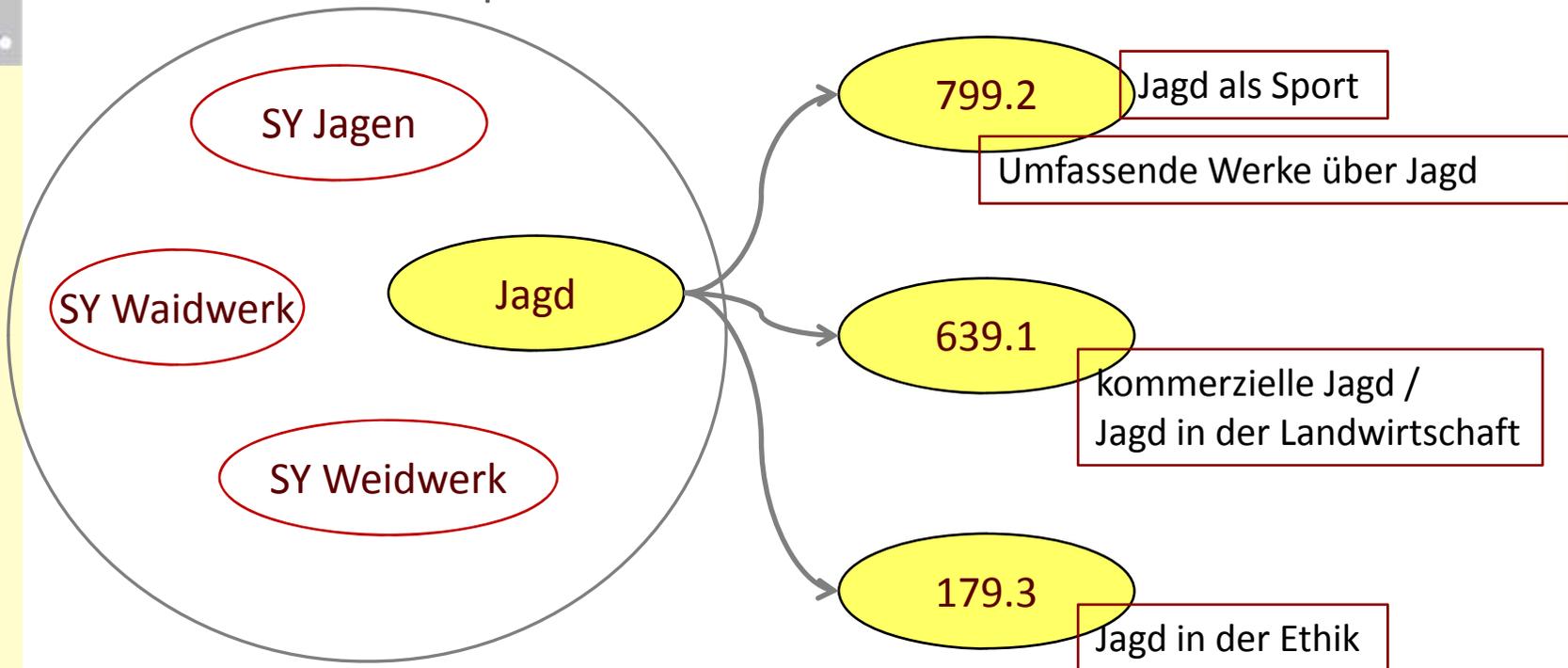
Fachhochschule Köln  
Cologne University of Applied Sciences

# CrissCross — Begriffliche Interoperabilität

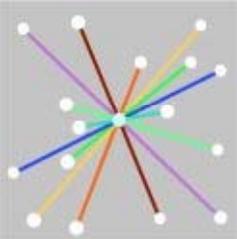


- unidirektional: SWD → DDC
- umfassend und spezifisch

Begriffliche Suche und Exploration  
in heterogenen Informationsräumen

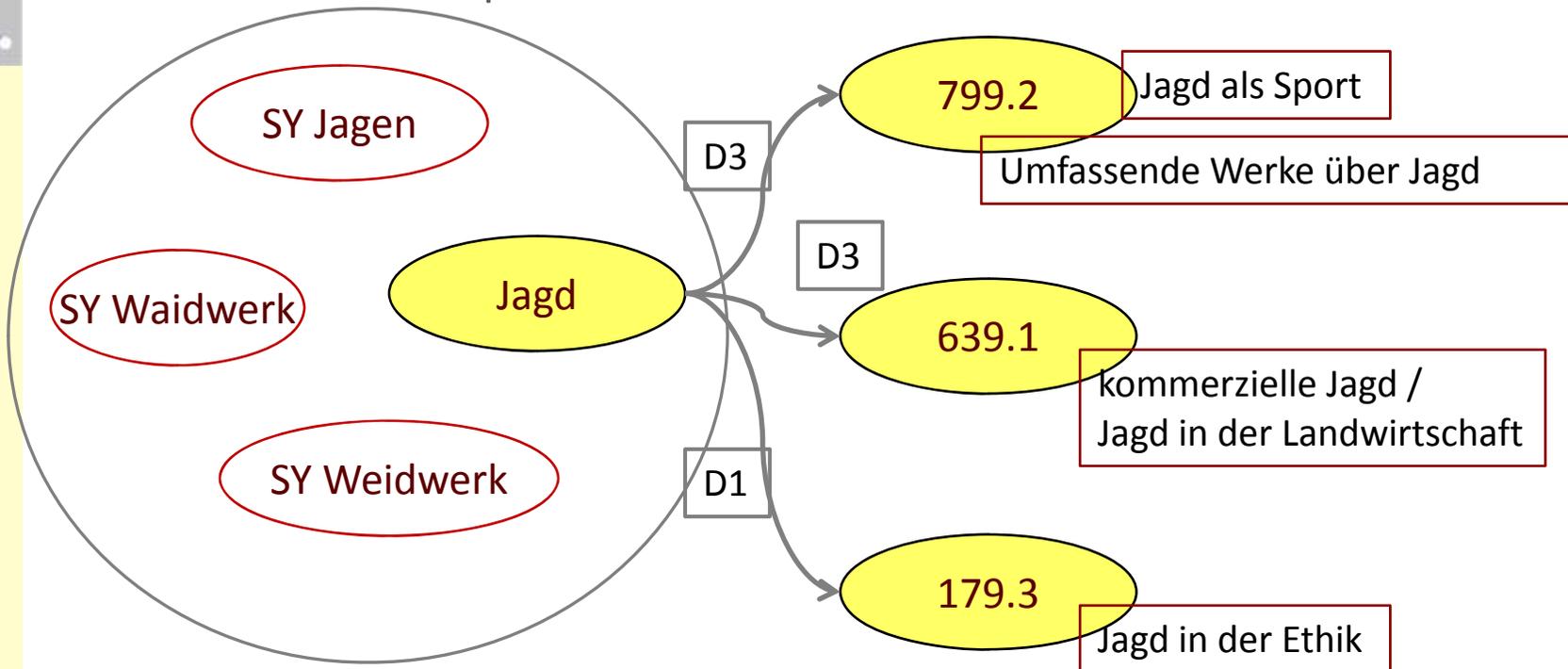


# CrissCross — Begriffliche Interoperabilität



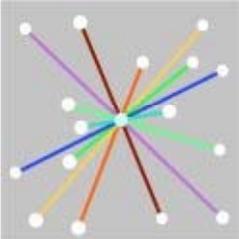
- unidirektional: SWD → DDC
- umfassend und spezifisch

Begriffliche Suche und Exploration  
in heterogenen Informationsräumen



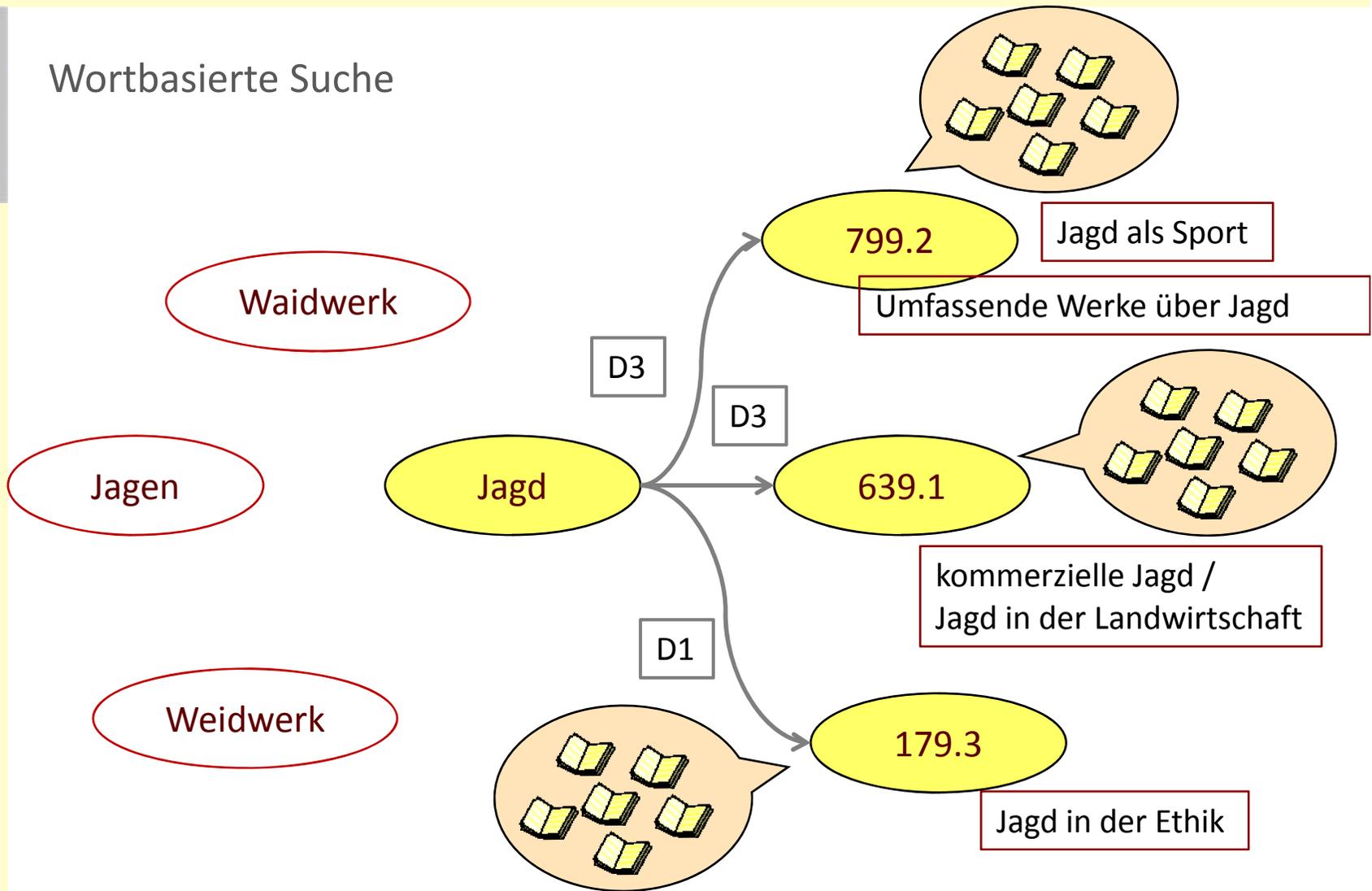
- differenziert

# CrissCross — Wortbasierte / Begriffliche Recherche

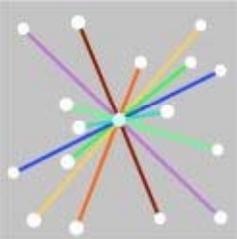


Wortbasierte Suche

Begriffliche Suche und Exploration  
in heterogenen Informationsräumen

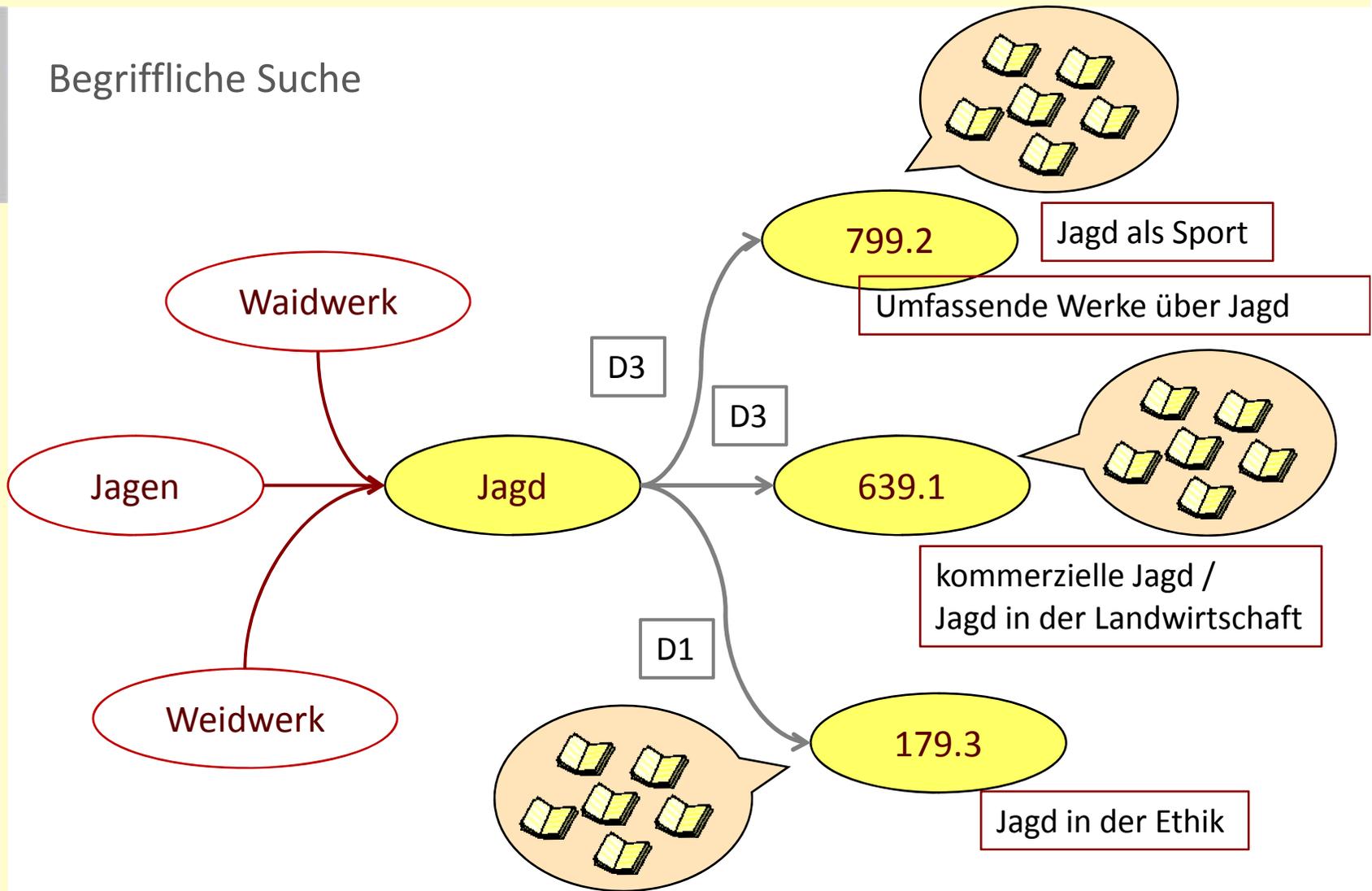


# CrissCross — Wortbasierte / Begriffliche Recherche

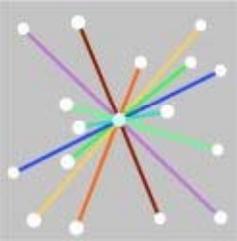


## Begriffliche Suche

Begriffliche Suche und Exploration  
in heterogenen Informationsräumen



# CrissCross — Differenzierte Interoperabilität



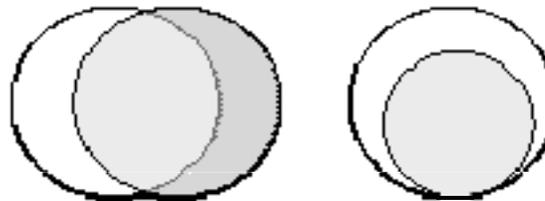
Begriffliche Suche und Exploration  
in heterogenen Informationsräumen

## 4 Determiniertheitsgrade

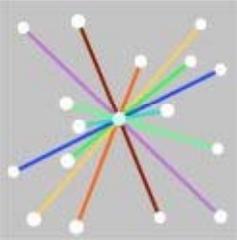
- zur Abbildung der Art der Beziehung zwischen SW u. DDC-Klasse
- Berücksichtigung der Mappingstrategien
- Orientierung an Thema-Klasse-Beziehungen der DDC

### Determiniertheitsgrad 3 = Wesentliche Übereinstimmung in CrissCross

- a) Der Bedeutungsumfang des SWs umfasst den (disziplinären) Aspekt der DDC Klasse und
- b) das durch das Schlagwort zum Ausdruck gebrachte Thema ist – unter dem (disziplinären) Aspekt betrachtet –
  - koextensiv mit DDC-Klasse
  - deckt mehr als die Hälfte des Inhalts der DDC-Klassenbenennung ab



# CrissCross — Differenzierte Interoperabilität



Begriffliche Suche und Exploration  
in heterogenen Informationsräumen

## 4 Determiniertheitsgrade

- zur Abbildung der Art der Beziehung zwischen SW u. DDC-Klasse
- Berücksichtigung der Mappingstrategien
- Orientierung an Thema-Klasse-Beziehungen der DDC

### Determiniertheitsgrad 3 = Wesentliche Übereinstimmung in CrissCross

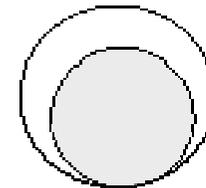
- a) Der Bedeutungsumfang des SWs umfasst den (disziplinären) Aspekt der DDC Klasse und
- b) das durch das Schlagwort zum Ausdruck gebrachte Thema ist – unter dem (disziplinären) Aspekt betrachtet –

- koextensiv
- deckt mehr als die Hälfte des Inhalts

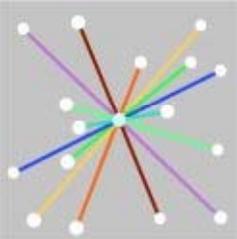
### Wesentliche Übereinstimmung in DDC

- Thema ist koextensiv mit Klasse
- Thema deckt mehr als die Hälfte des Inhalts der Klassenbenennung ab

- Standardschlüssel dürfen angehängt werden



# CrissCross — Thematische Suche



Begriffliche Suche und Exploration  
in heterogenen Informationsräumen

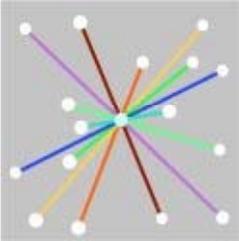
## Grundlage

Dokumentspezifische Begriffe, die den inhaltlichen Gegenstand eines Dokuments beschreiben

Thematische Suchen in heterogenen Informationsräumen erfordern eigentlich *thematische Interoperabilität*. Begrenzt kann indes auch auf der Basis von begrifflicher Interoperabilität eine thematische Recherche i.e.S. durchgeführt werden. Dies setzt voraus, dass

- die **Relation zwischen Begriffen in der jeweiligen Dokumentations-sprache und dokumentspezifischen Begriffen** klar umrissen werden kann
- die **Anwendung** der spezifischen Dokumentations-sprachen , wie sie in den jeweiligen Regeln festgelegt sind, bei der Gestaltung der Such-funktionalitäten mitberücksichtigt wird

# CrissCross — Thematische Suche



Begriffliche Suche und Exploration  
in heterogenen Informationsräumen

## Schlagwortketten nach RSWK / postkoordinierende Indexierung

Jagd  
Jagd / Brauchtum / Österreich  
Jagd / Bedrohte Tiere / Nachhaltige Entwicklung / Lateinamerika/  
Jagd / Tourismusindustrie / Auslandsreise  
Jagd / Wildtiere / Rotbuche / Waldbestand / Ökologisches Gleichgewicht / Deutschland

?

SWD Jagd

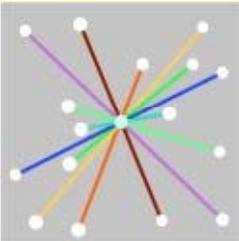
DDC 799.2

## Dokumentspezifische DDC

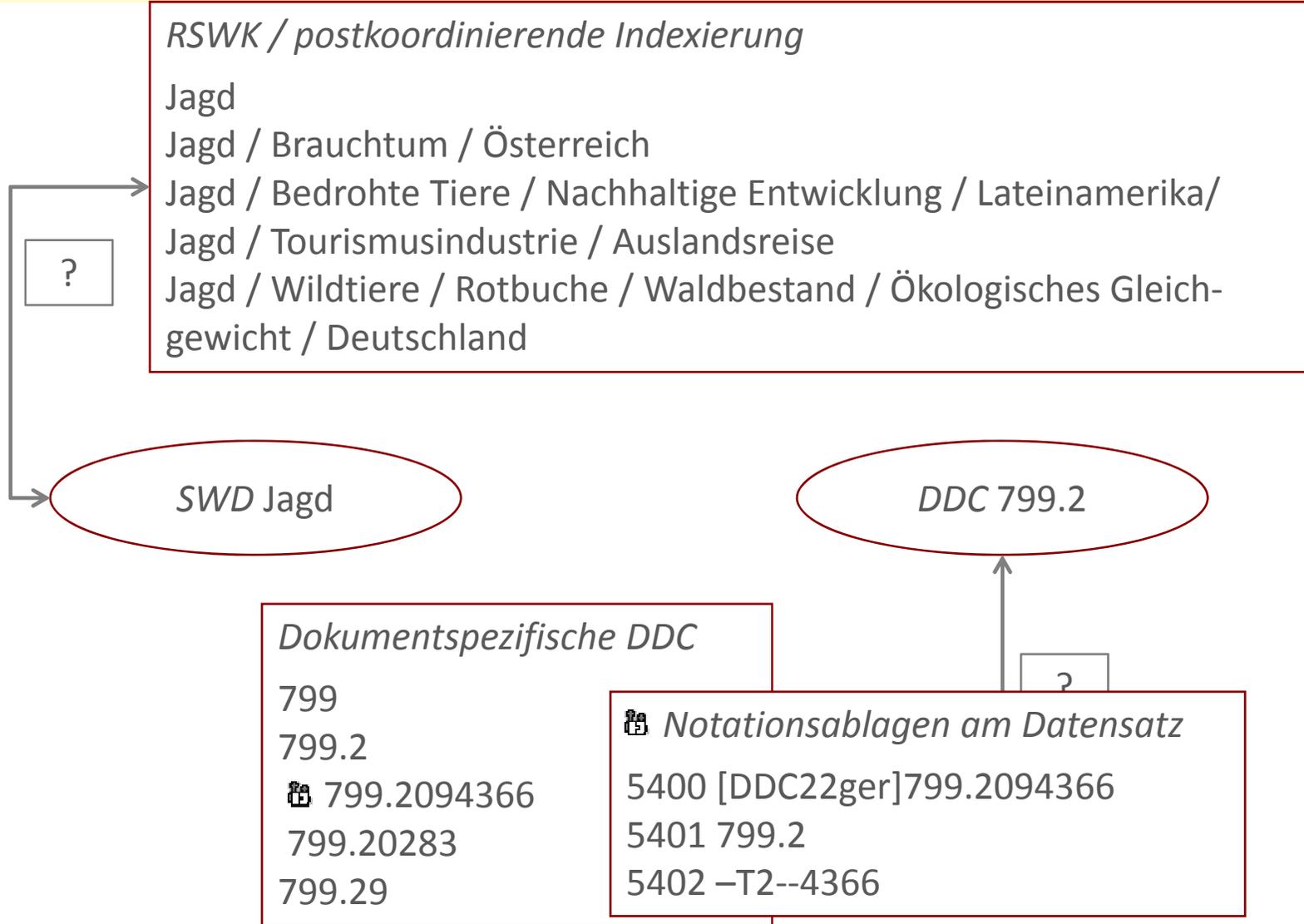
799  
799.2  
📖 799.2094366  
799.20283  
799.29

?

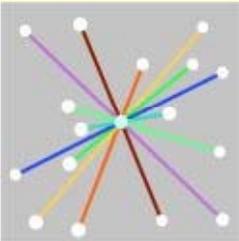
# CrissCross — Thematische Suche



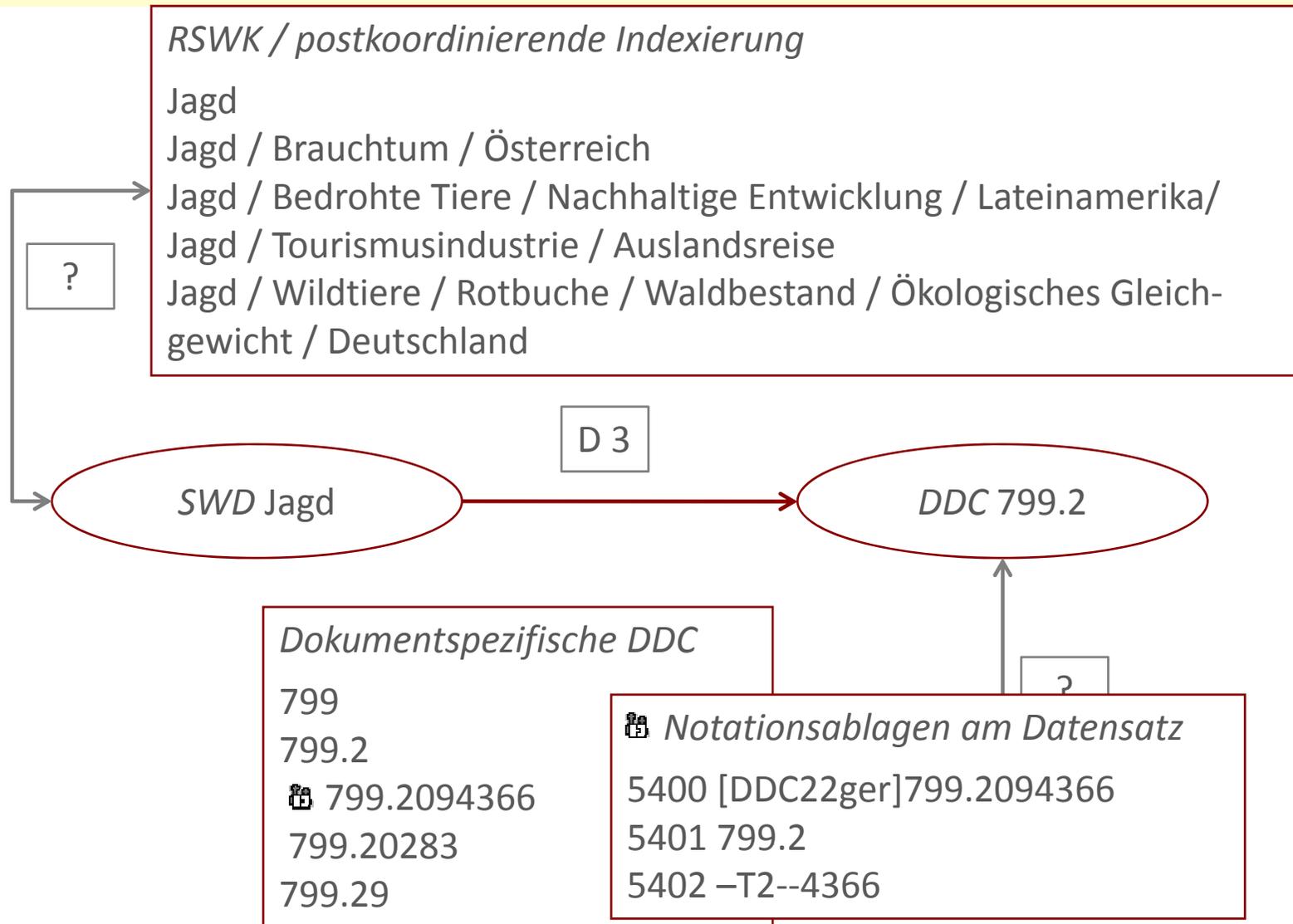
Begriffliche Suche und Exploration  
in heterogenen Informationsräumen



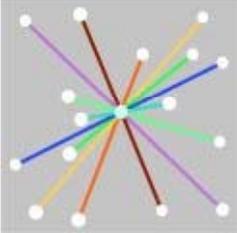
# CrissCross — Thematische Suche



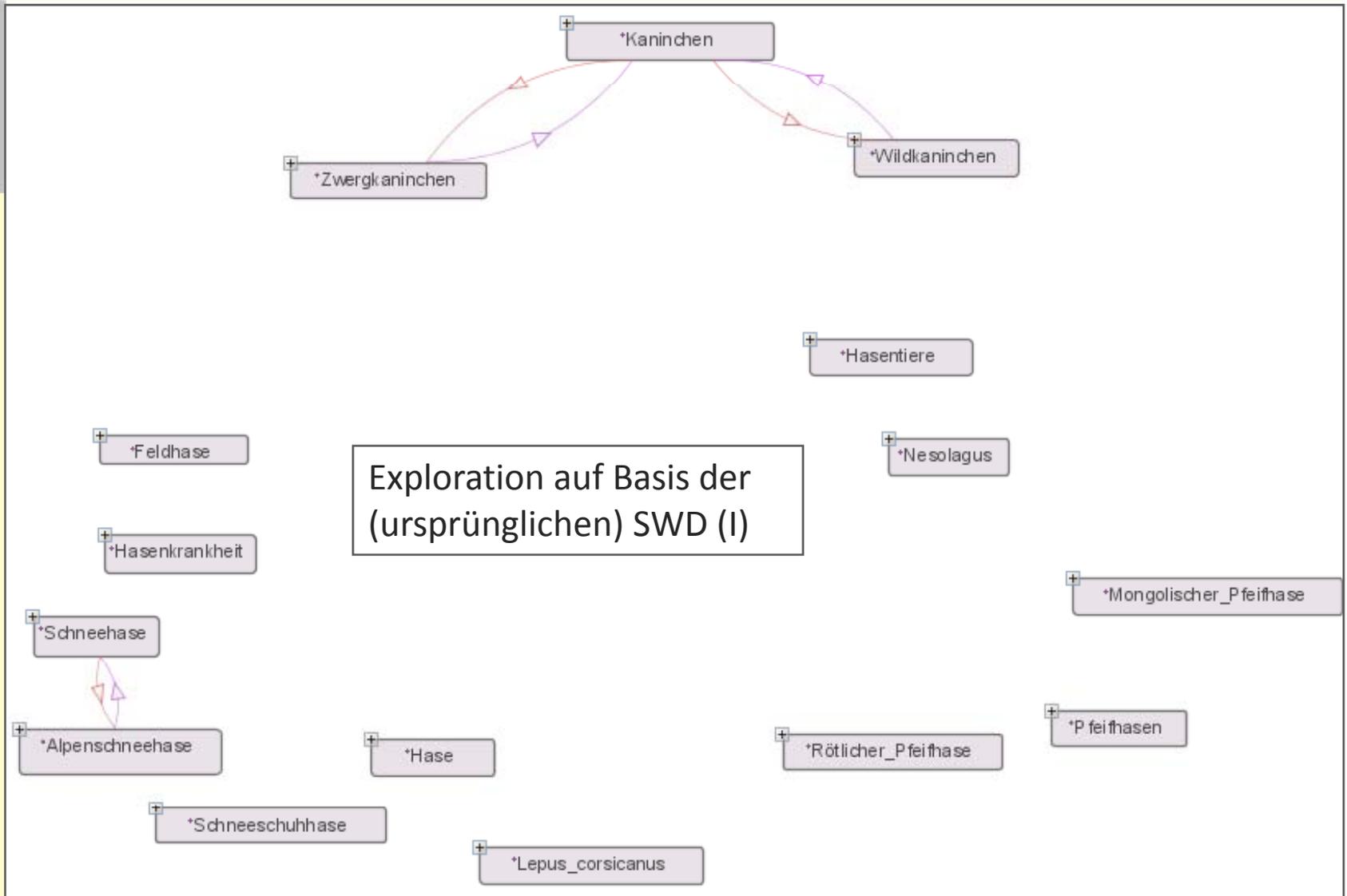
Begriffliche Suche und Exploration  
in heterogenen Informationsräumen



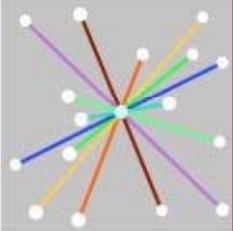
# CrissCross — Wissensexploration



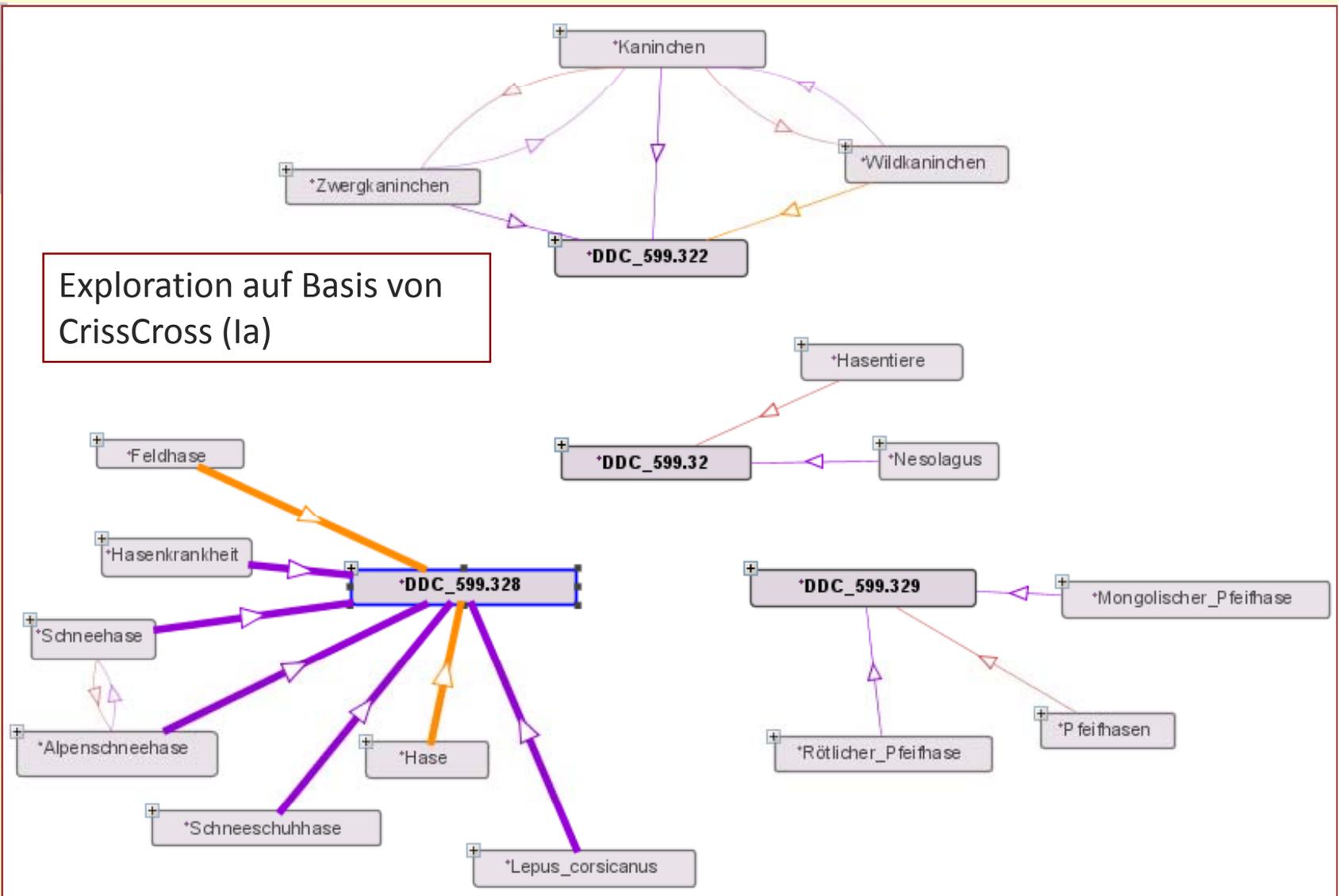
Begriffliche Suche und Exploration  
in heterogenen Informationsräumen



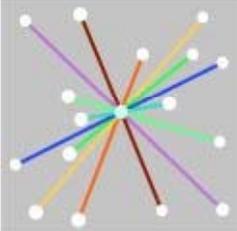
# CrissCross — Wissensexploration



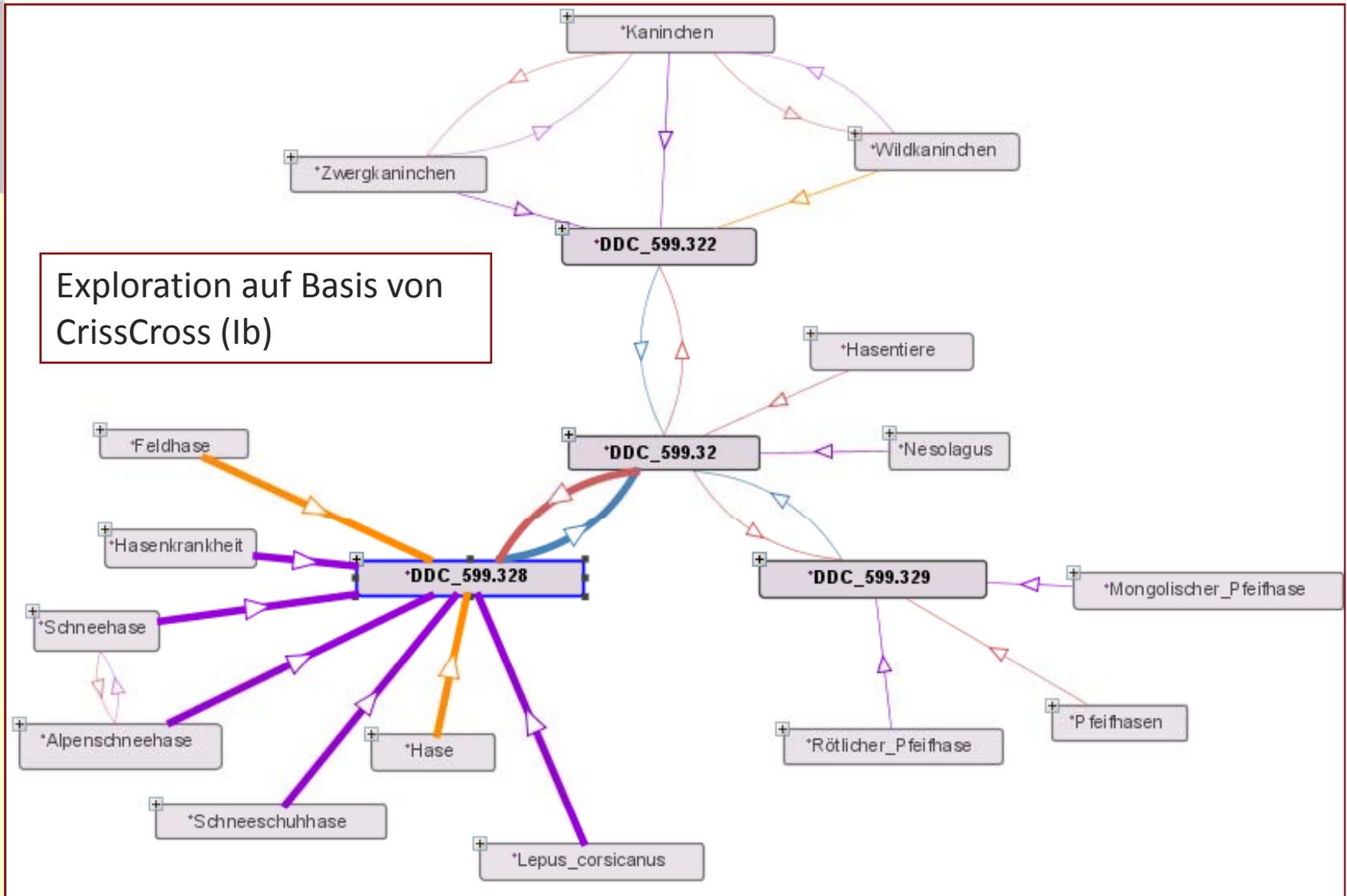
Begriffliche Suche und Exploration  
in heterogenen Informationsräumen



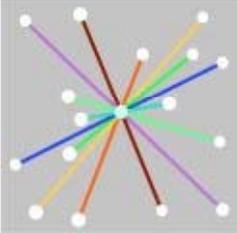
# CrissCross — Wissensexploration



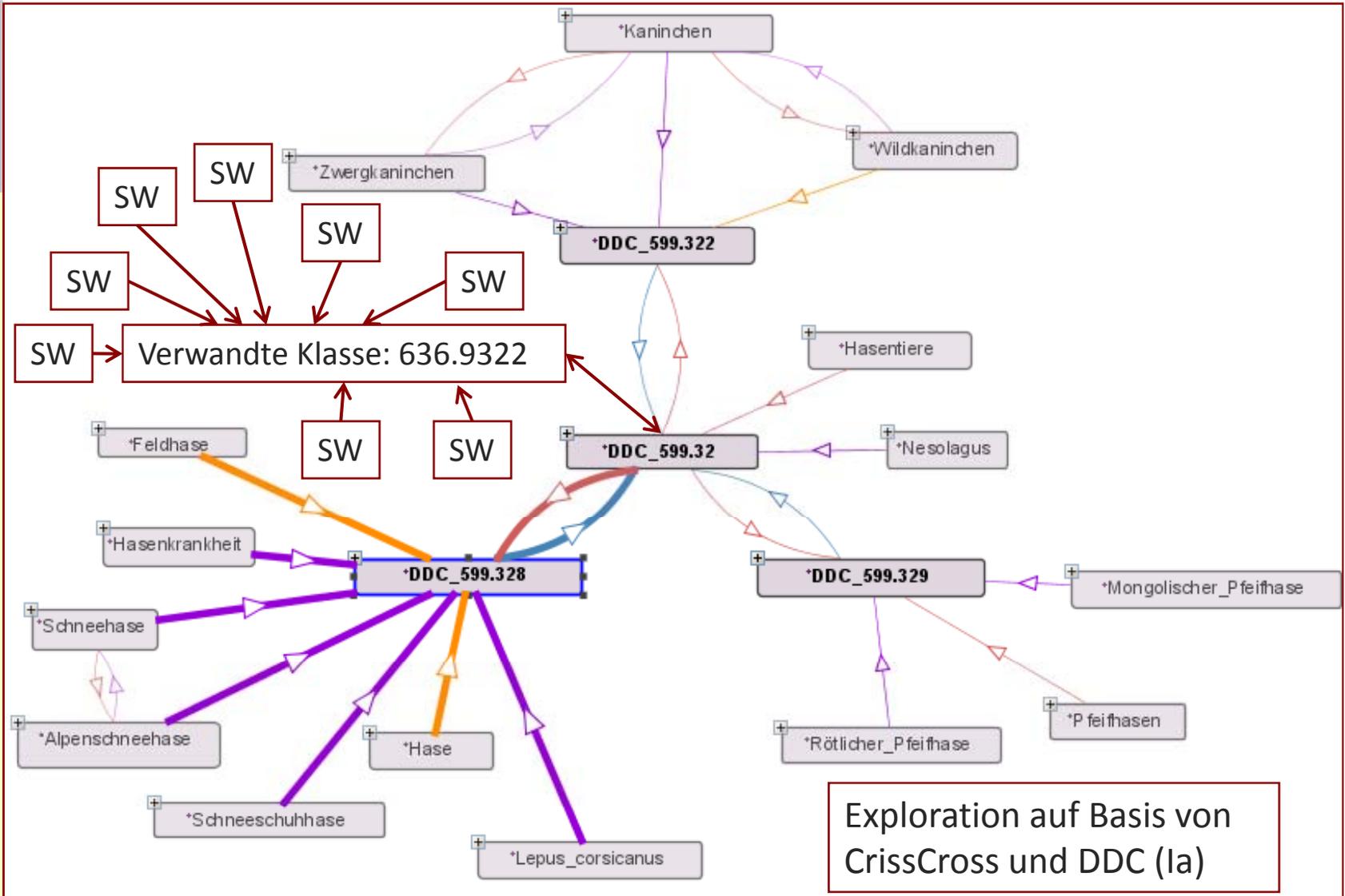
Begriffliche Suche und Exploration  
in heterogenen Informationsräumen



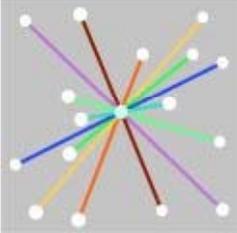
# CrissCross — Wissensexploration



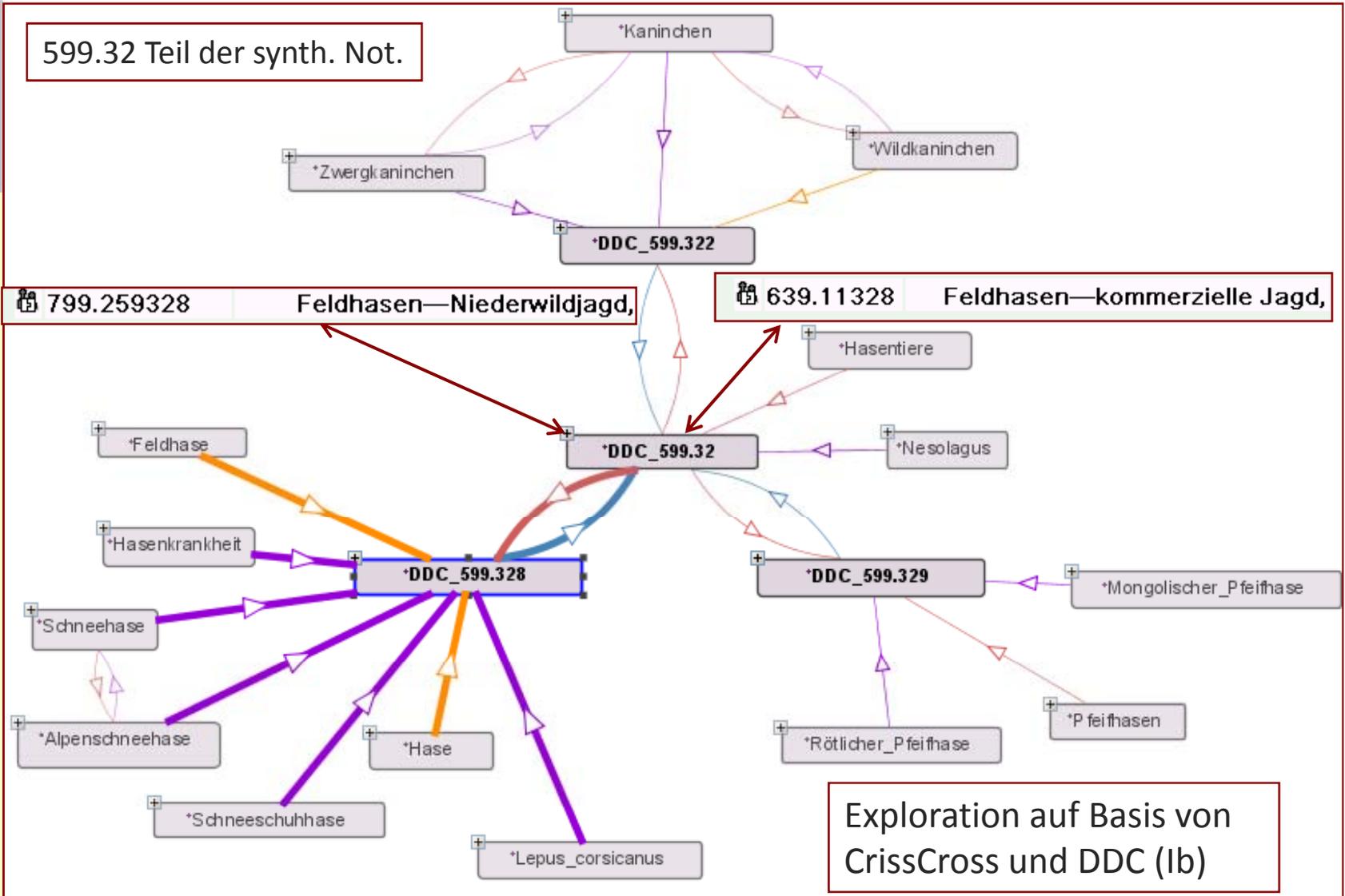
Begriffliche Suche und Exploration  
in heterogenen Informationsräumen



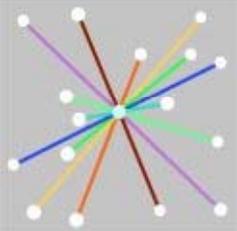
# CrissCross — Wissensexploration



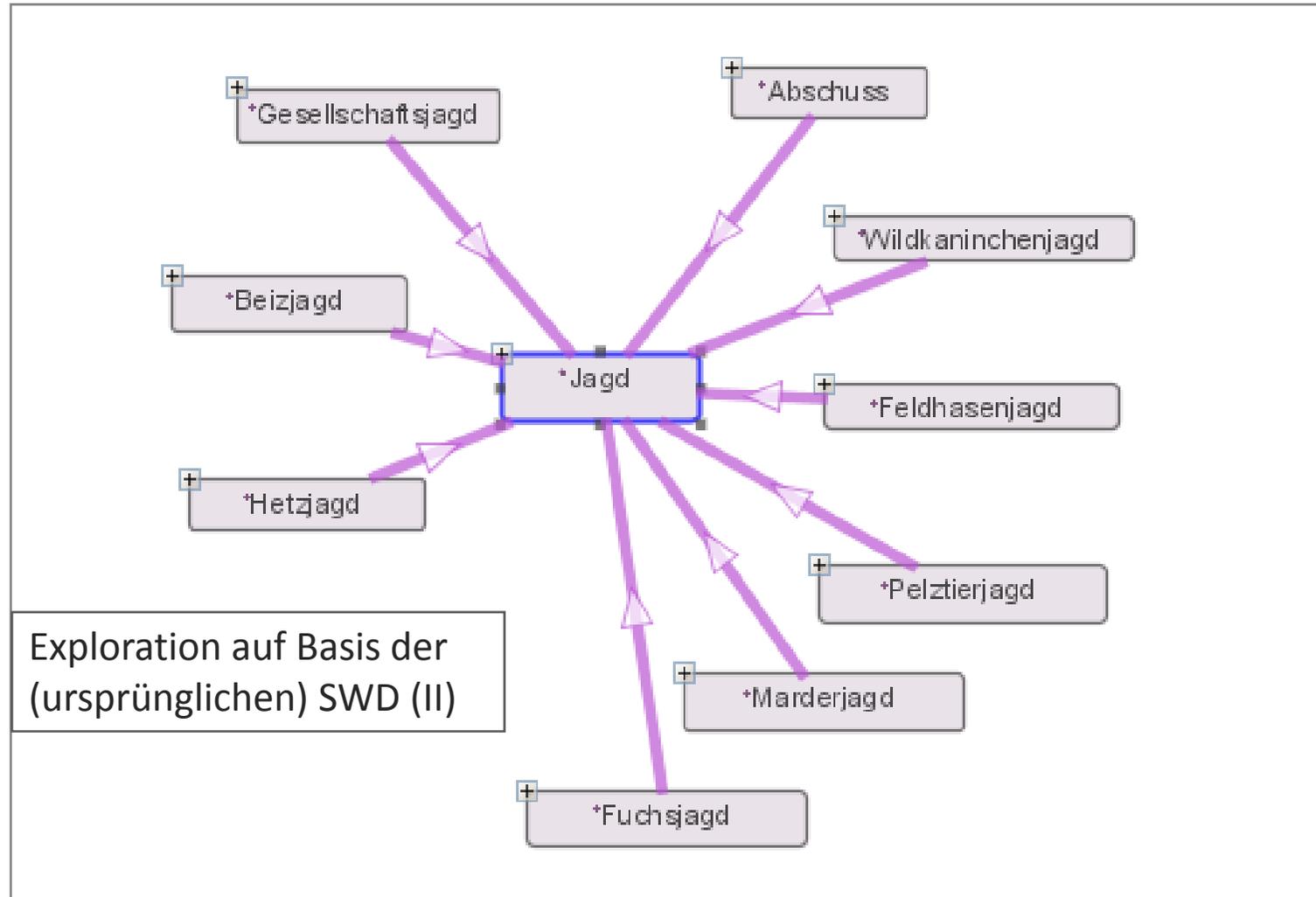
Begriffliche Suche und Exploration  
in heterogenen Informationsräumen



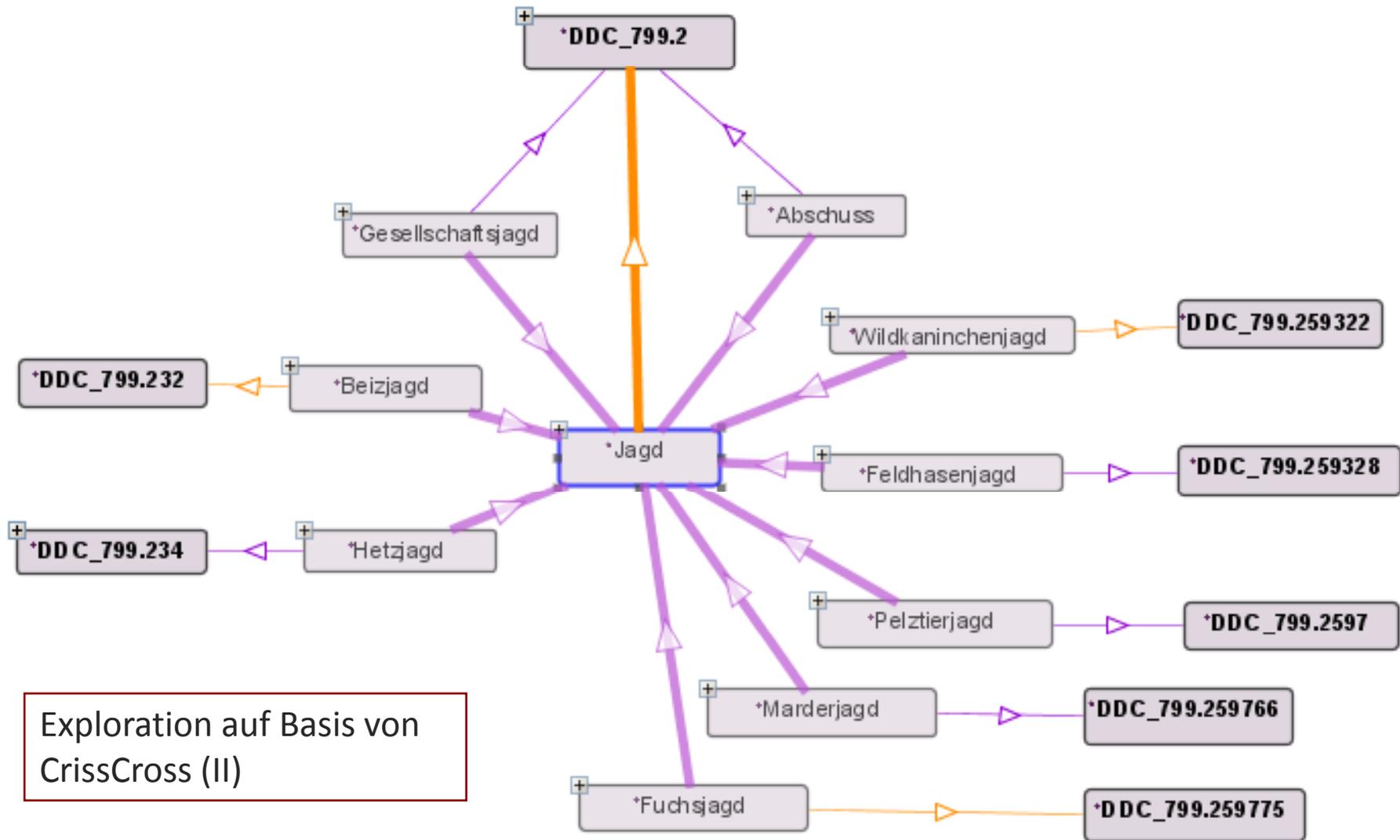
# CrissCross — Wissensexploration



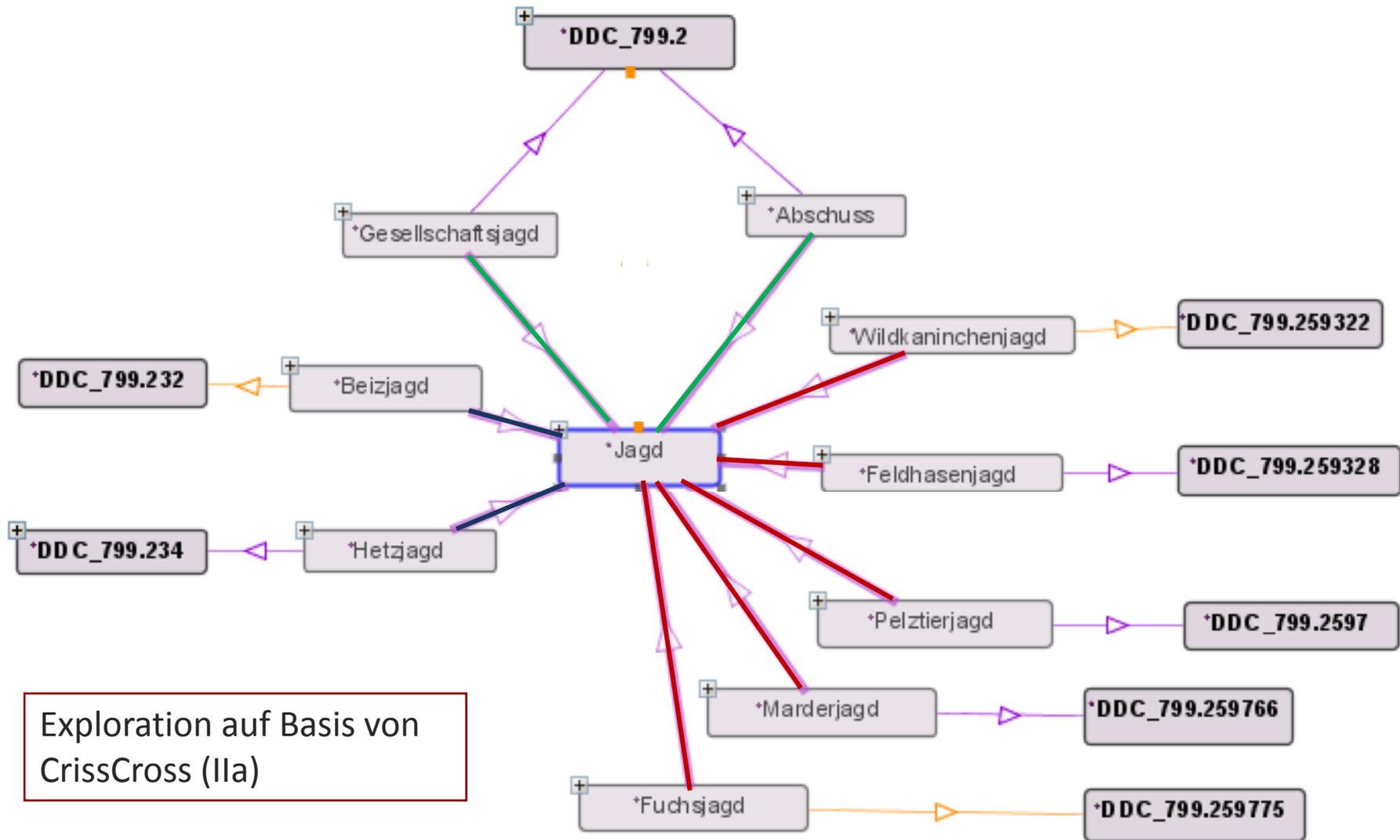
Begriffliche Suche und Exploration  
in heterogenen Informationsräumen



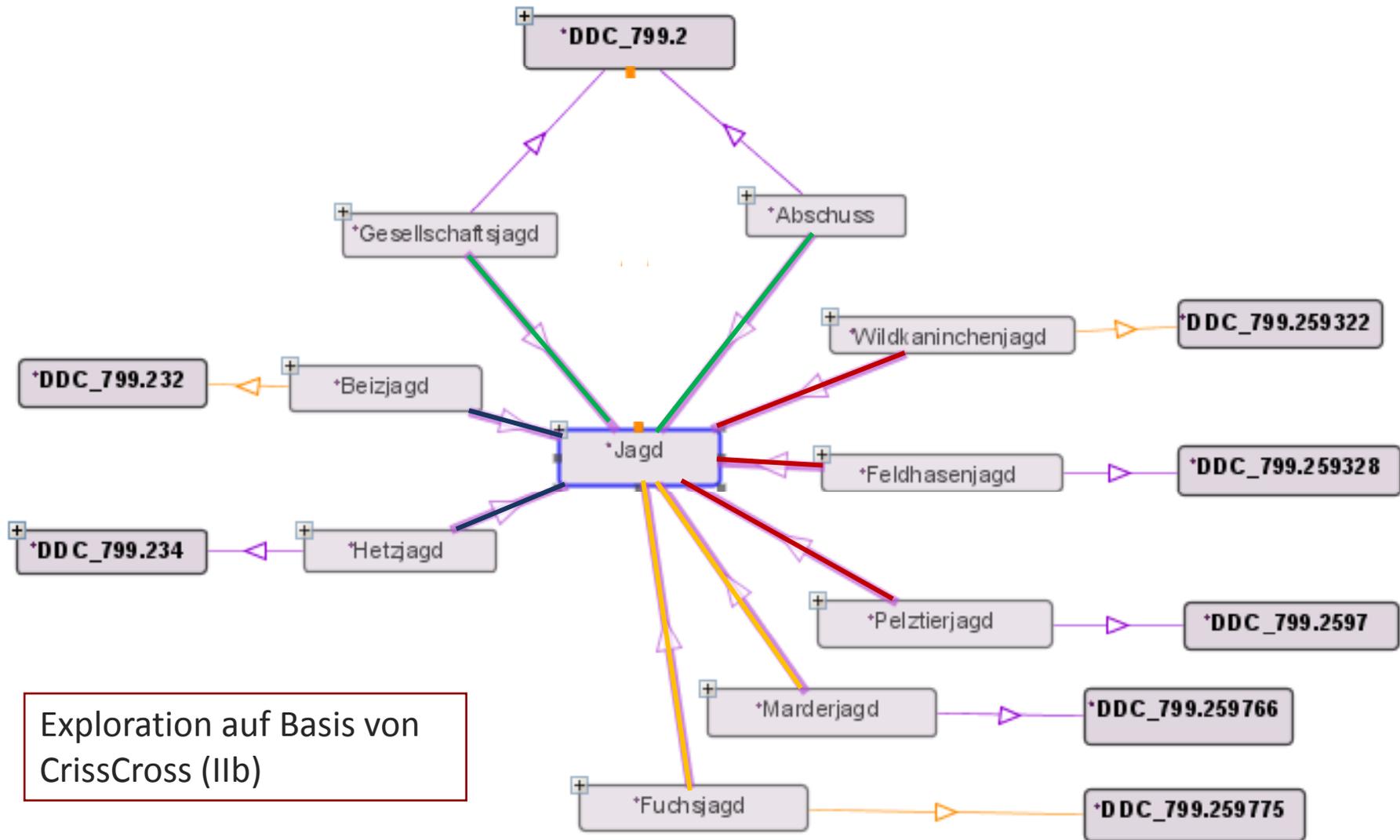
# CrissCross — Wissensexploration



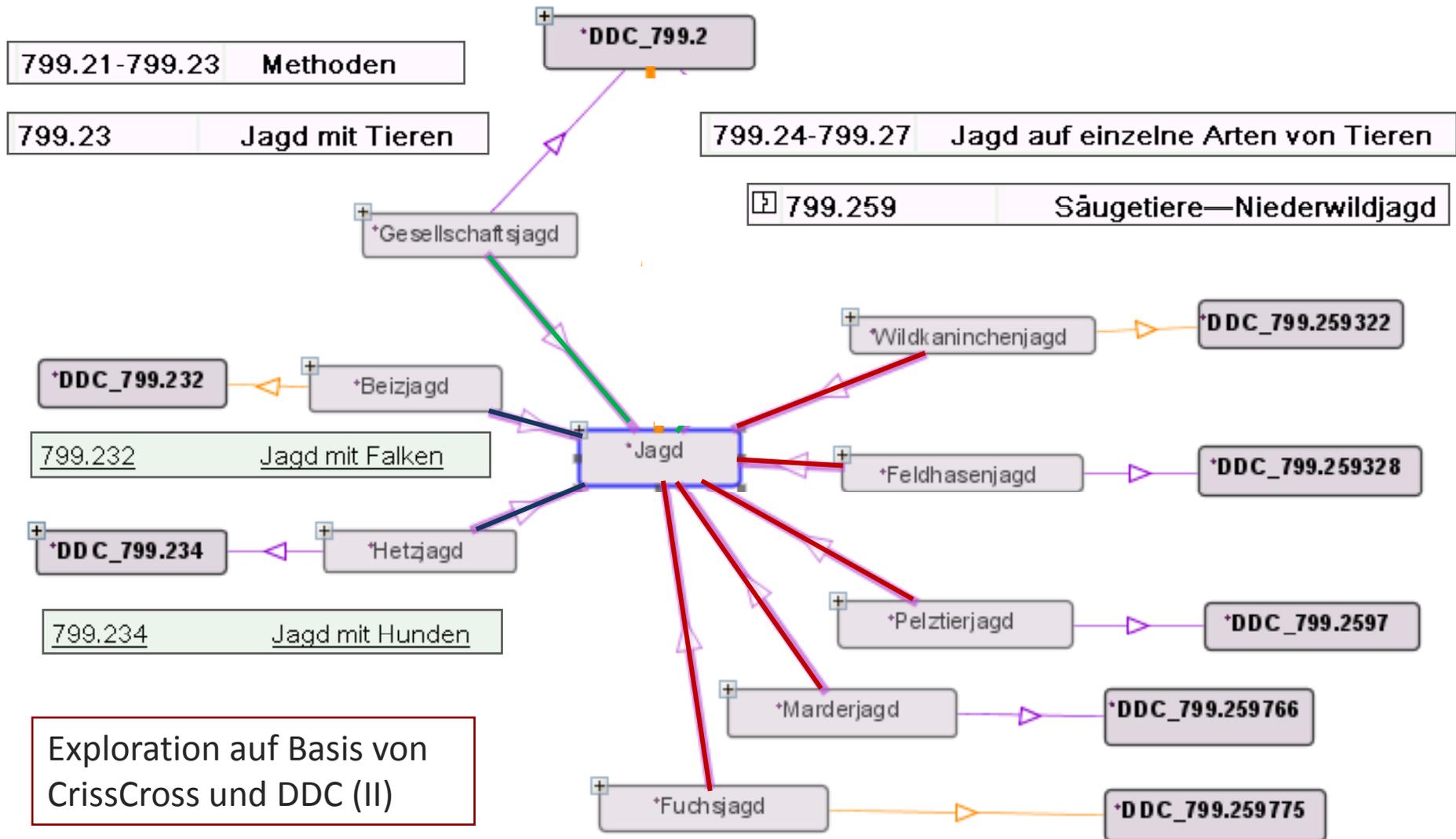
# CrissCross — Wissensexploration



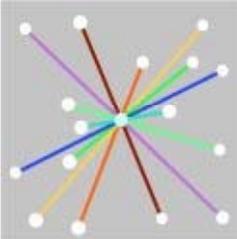
# CrissCross — Wissensexploration



# CrissCross — Wissensexploration



# CrissCross — Strukturierung von Treffermengen

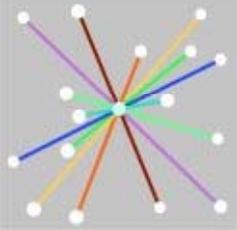


## (I) Ranking von Literatur, die mit einer DDC-Notation erschlossen wurde

Begriffliche Suche und Exploration  
in heterogenen Informationsräumen

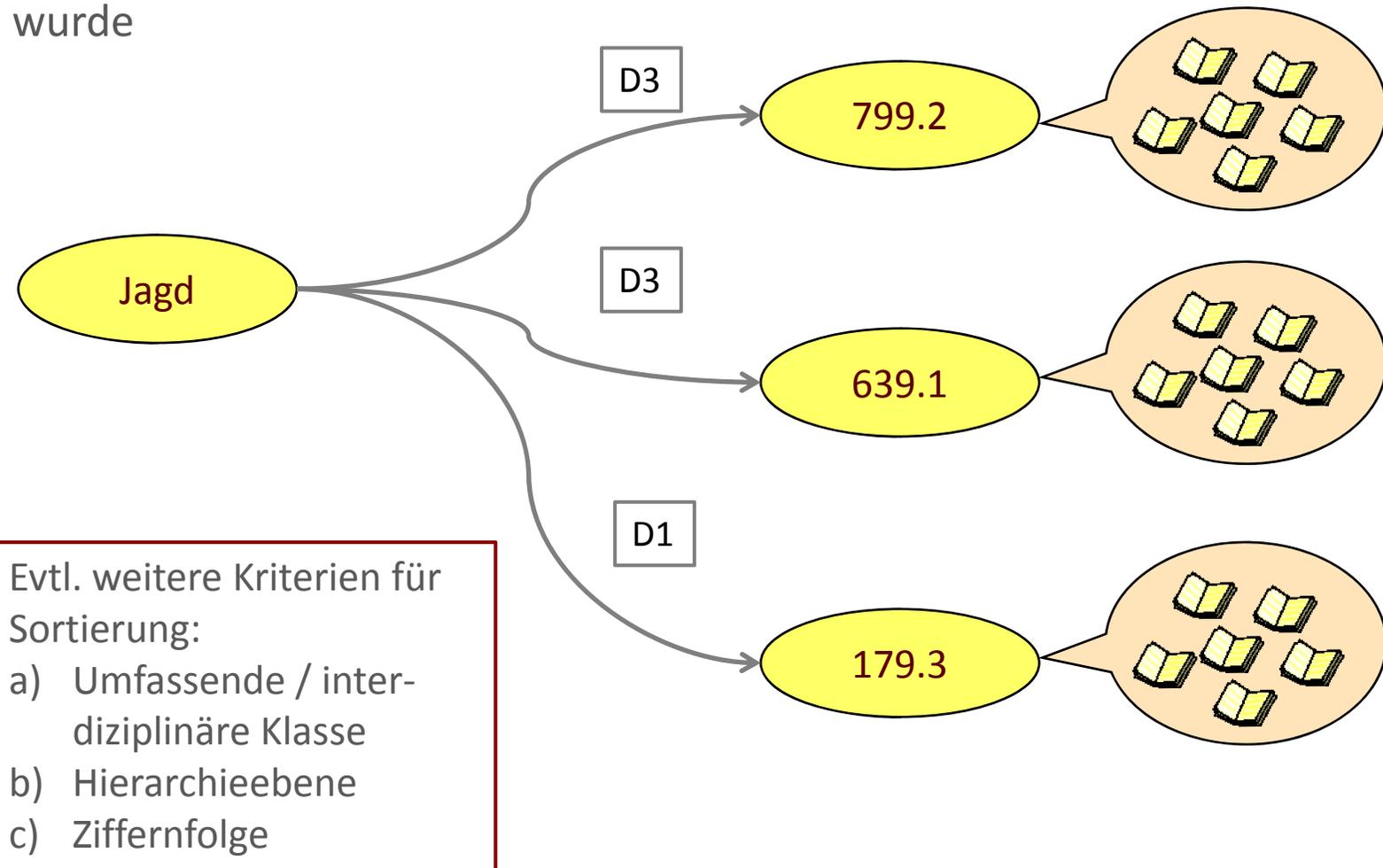
| DDC-Notation  | Thema   | SWD-Schlagwörter  | Relevanz      |
|---------------|---|---|---------------|
| 790           | <u>Sport, Spiele, Unterhaltung</u>                                |   |               |
| 799           | <u>Fischfang, Jagd, Schießen</u>                                  | s Fallenjagd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

# CrissCross — Strukturierung von Treffermengen

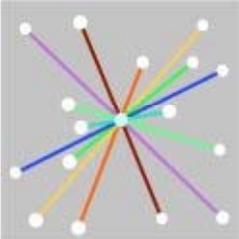


Begriffliche Suche und Exploration  
in heterogenen Informationsräumen

(II) Ranking von Literatur, die mit mehreren DDC-Notationen erschlossen wurde

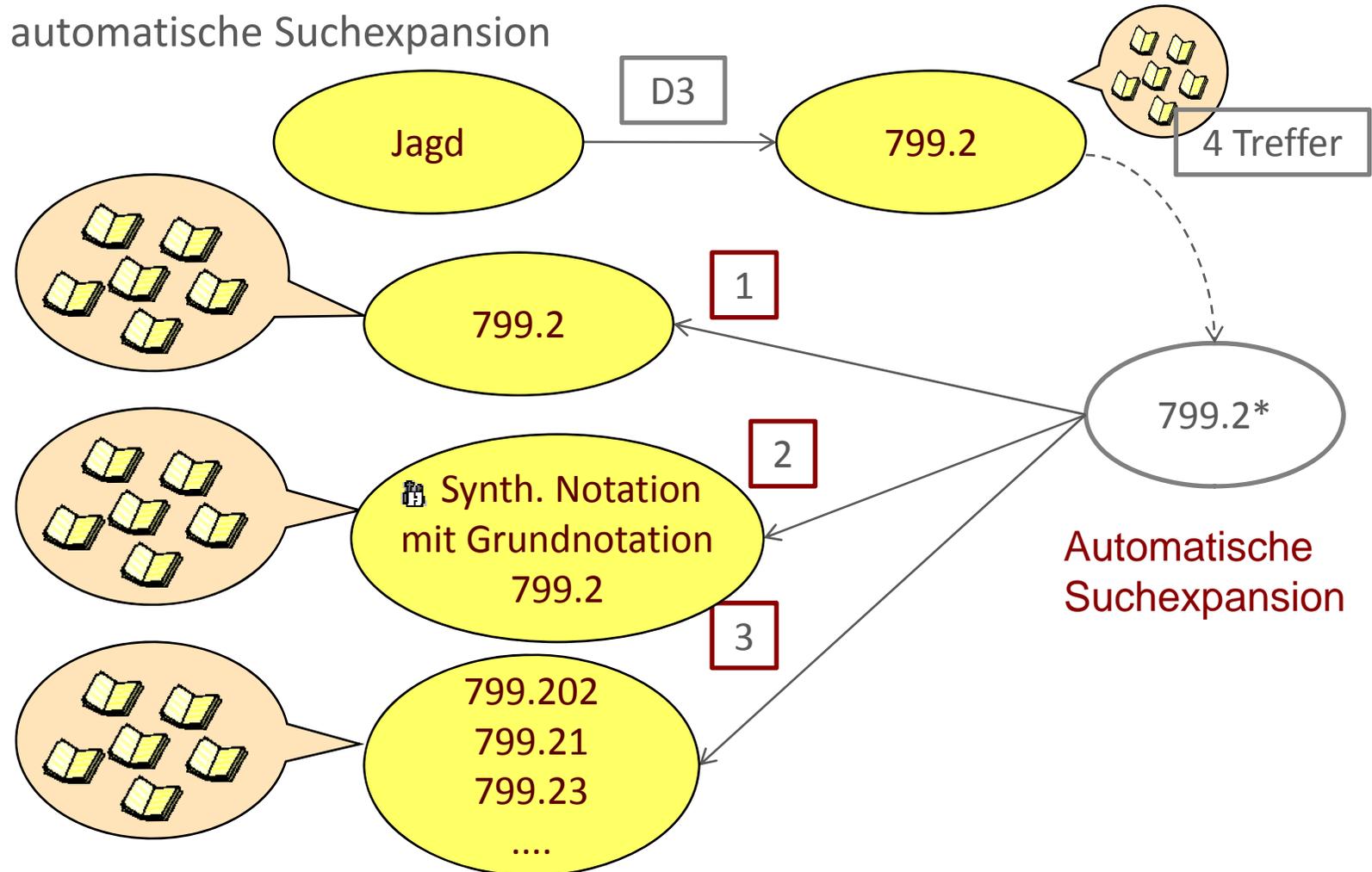


# CrissCross — Strukturierung von Treffermengen (II)

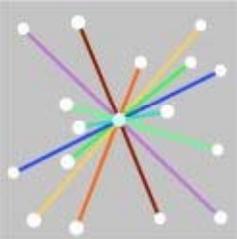


Begriffliche Suche und Exploration  
in heterogenen Informationsräumen

(III) Ranking von DDC-erschlossener Literatur im Anschluss an eine automatische Suchexpansion



# Zusammenfassung



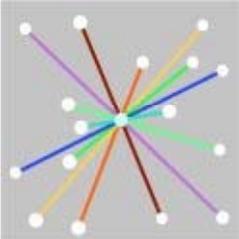
Begriffliche Suche und Exploration  
in heterogenen Informationsräumen

Die theoretischen Reflexionen zu CrissCross brachten eine Präzisierung hinsichtlich

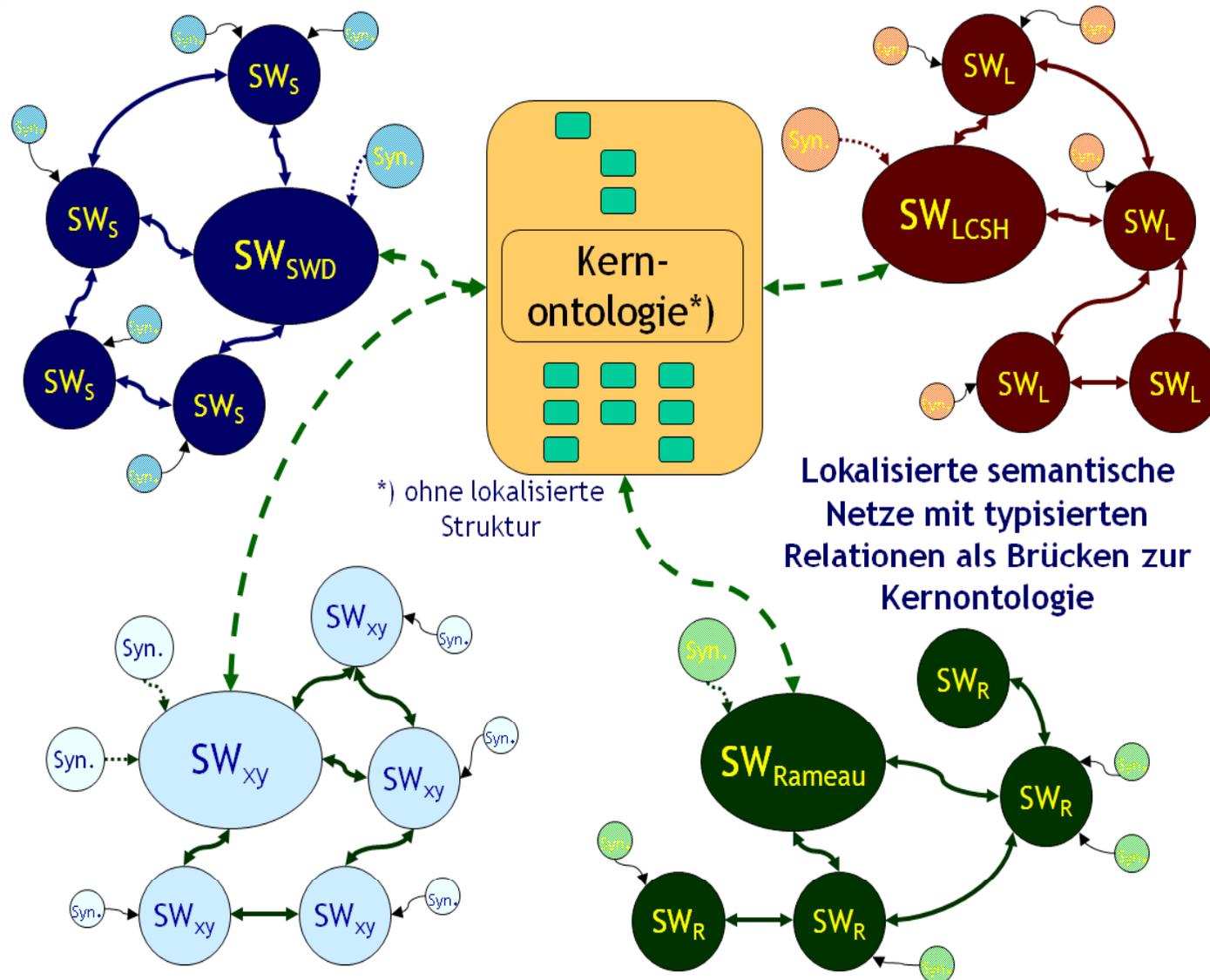
- Unterscheidung von Art/en der semantischen Interoperabilität
- Möglichkeiten zur Gestaltung von Recherchefunktionalitäten
  - a) unter Berücksichtigung der erstellten semantischen Interoperabilität und
  - b) in Hinblick auf *Recall* und *Precision*
- Strukturierung von Treffermengen / Ranking nach Relevanz auf der Basis
  - a) der erstellten semantischen Interoperabilität und
  - b) der gegebenen Funktionalitäten des Zielvokabulars

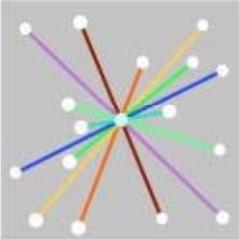
Die erarbeiteten Grundlagen können auch für die Entwicklung umfassenderer Interoperabilitätsmodelle als Ausgangspunkt dienen.

# Zukunftsvision?



Begriffliche Suche und Exploration  
in heterogenen Informationsräumen





Begriffliche Suche und Exploration  
in heterogenen Informationsräumen

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

