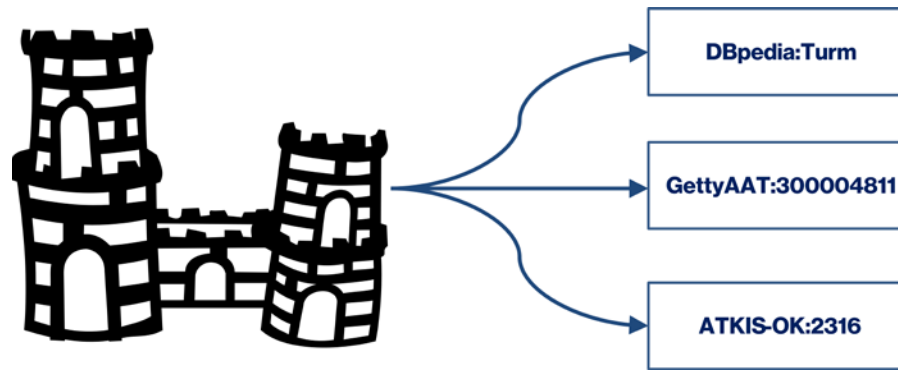


Thesauri-Vernetzung dank «Labeling System» – ein Use case anhand Schweizer Rechtsquellen



Pascale Sutter, Rechtsquellenstiftung des
Schweizerischen Juristenvereins

Open-Access-Tage, 07./08. September 2015, Zürich

Session 1: (Linked) Open Data, Thesauri und kulturelles Erbe

Dank

Prof. Dr. [Kai-Christian Bruhn](#) und [Florian Thiery](#),
i3mainz – Institut für raumbezogene Informations-
und Messtechnik an der Hochschule Mainz

[Giovanni Colavizza](#), DHLAB der ETH Lausanne

Dr. [Michael Piotrowski](#), Leibniz-Institut für
Europäische Geschichte, Mainz

[Jonas Schneider](#), HLS

Aufbau

- Ziel von hist-metadata.net
- Was kann/ist das Labeling System (LS)?
- Begriffsdefinitionen
- Der «Labeling Approach»
- LS-Prädikatenkanon
- Verortung der Vokabulare
- Beispiel: Vornamen-Normdatei
- Ergebnisse der Probephase

Ziel von hist-metadata.net

Mehrsprachige kontrollierte Vokabulare und Normdateien für die Geisteswissenschaften als Linked Open Data (LOD) zur Verfügung zu stellen, die den hohen wissenschaftlichen Qualitätsanforderungen dank einheitlichen Standards genügen.



Was kann/ist das LS? (1)

- Freie Webanwendung zur Erstellung eines mehrsprachigen kontrollierten Vokabulars basierend auf LOD und SKOS (projektspezifisches Vokabular)
Resultate sind frei verfügbar (Vokabulare und Labels unter CC-BY-Lizenz)
- Nutzung des «Labeling Approach»: Anreicherung eines Vokabular-Begriffs mit Links zu Ressourcen und Konzepten in der Linked Data Cloud, z. B. Definitionen der Begriffe oder Funktionen

Was kann/ist das LS? (2)

- Erstellen von Hierarchien und Relationen innerhalb des Vokabulars und zu externen Ressourcen und damit die Möglichkeit zur Implementierung einer eigenen Ontologie im LS



Begriffsdefinitionen

Term: Begriff mit natürlichsprachlicher Definition

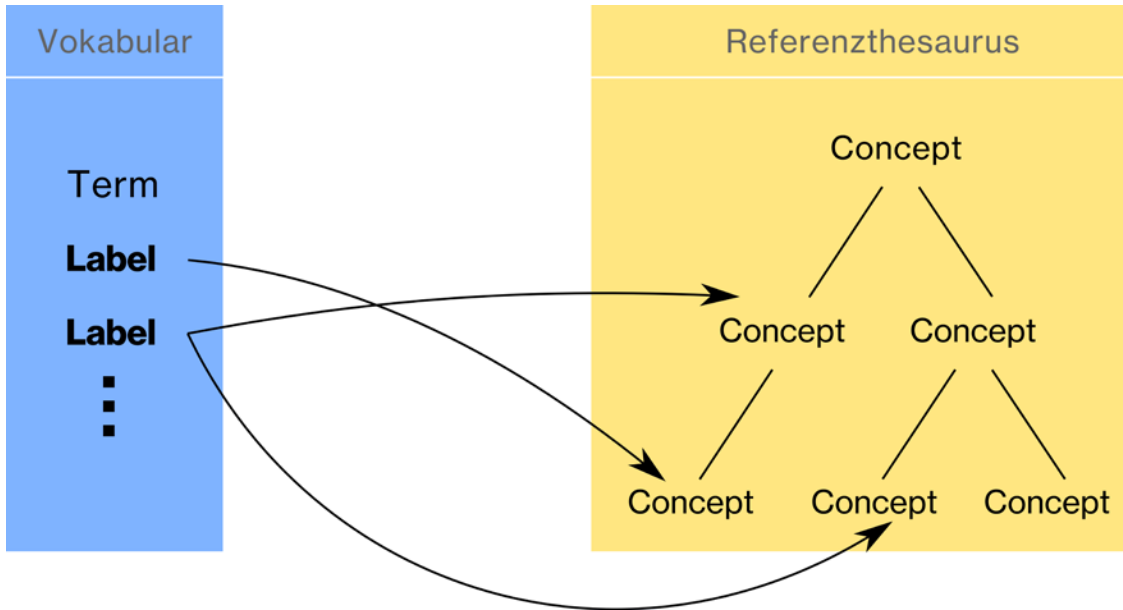
Label: Begriff mit natürlichsprachlicher Definition UND Link zu einem oder mehreren Konzepten in einem Referenzthesaurus

Concept: HTTP adressierbare Ressource in der Linked Data Cloud, die einen Begriff näher beschreibt (SKOS, HTML, OWL etc.) und von anderen im Semantic Web genutzt wird

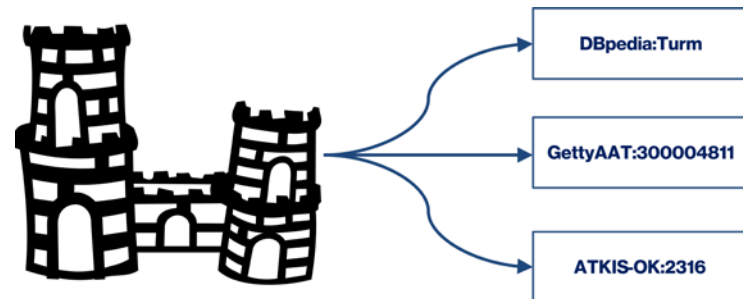
Vokabular: Sammlung von mehrsprachigen Terms und Labels

Referenzthesaurus: Sammlung von HTTP adressierbaren Konzepten, z. B. Getty AAT oder Wikidata

Der «Labeling Approach»



Graphiken: Florian Thiery

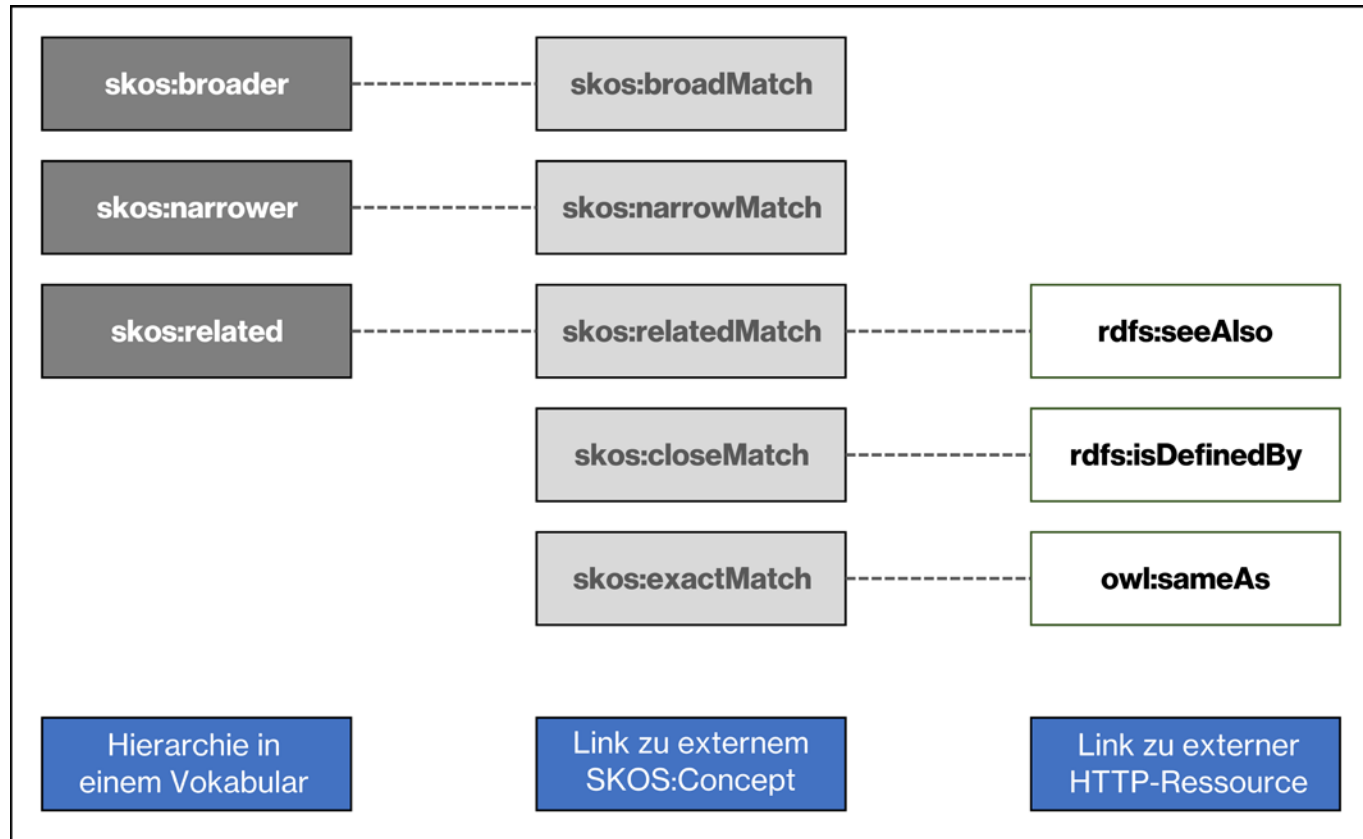


LS-Prädikatenkanon

- Das LS lässt nur bestimmte Prädikate zur Verlinkung von Labels untereinander zu.
- Durch geschicktes Anwenden dieser Prädikate kann jedoch eine eigene Ontologie geschaffen werden (vgl. Vornamen-Normdatei der SSRQ).

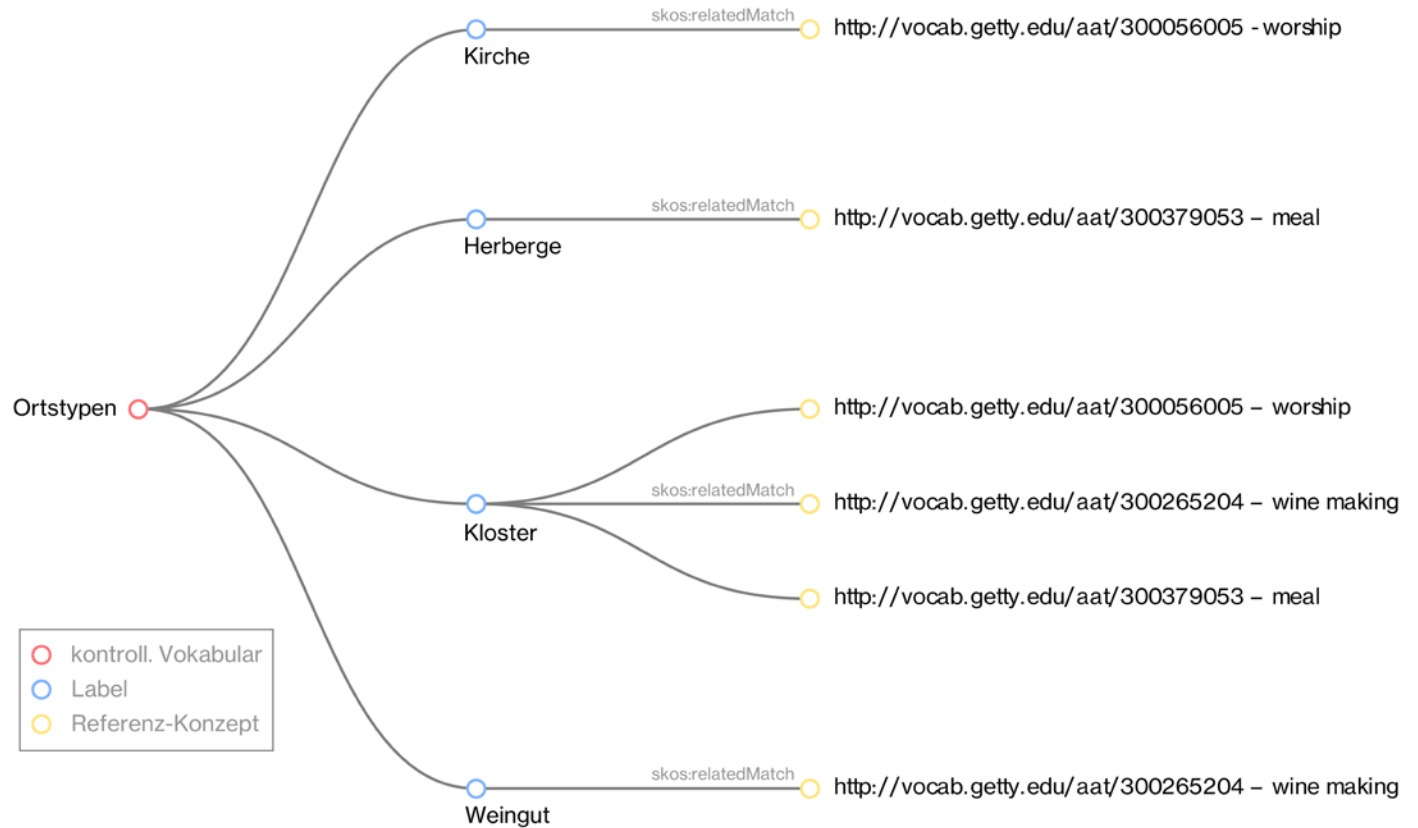


LS-Prädikatenkanon



Graphik: Florian Thiery

Beispiel: Ortsfunktionen



Graphik: Florian Thiery

Was kann/ist das LS? (3)

- basiert auf Standards, z. B. RDF, OWL, SKOS, Dublin Core und Open-Source-Komponenten
- Publikation des Vokabulars und der Labels als URI mittels RDF und SKOS und Verlinkung in die LOD Cloud
- Speicherung von Metadaten zum Erfassungs-, Änderungs- und Löschprozess
- Zusammenarbeit zwischen Forschern möglich, durch das Teilen der Inhalte mit anderen Nutzern
- Externer Zugriff mittels SPARQL Endpoint und REST Interface



Was kann das LS (noch) nicht?

- Das LS ist bislang nur ein Prototyp.

Applikation: <http://labeling.i3mainz.hs-mainz.de/>

- Es sind noch nicht alle Dinge der Ontologie umgesetzt. Eine intuitive Eingabe- und Administrationsoberfläche fehlt.
(Teil des Projekts histHub.net bzw. des Produkts hist-metadata.net)

Ontologie: <http://labeling.i3mainz.hs-mainz.de/ontology/>

Verortung des SSRQ-Vokabulars

SSRQ SG III/2, S. 1539 (Ausschnitt)

882⁵, 1048^{3,14}, 1114^{14,15}, 1134¹⁵, 1138⁴⁷, 1144³⁶, *s. auch* atzung, böveli, bovel, etzen, gäißwäide, grasmiet, küe weyden, mägeri, meyenweid, roßweiden, sale, sommerweide, stafel, stos, tagwaidlin, trib unnd trätt, waydenn, weid, *Weideordnung*, *Weiderecht*, *Weideschmalz*, weidgelt, weidung, wunn und weid, wunn

Weideordnung 577²⁹, 591³⁴

Weiderecht 4⁴¹, 243²¹, 296⁶, 391⁹, 511³¹, 532³⁶, 658⁸, 673^{5,9}, 896²⁶, 1048¹⁵, 1049³⁸, 1144³⁴, 1145¹², *s. auch* atzung, getraib, tratt recht, tratt, trattbefreyung, weidgang, weidsgerechtigkeit

Weideschmalz 321¹⁷, 400¹, *s. auch* weid-schmaltz

weidgang waidgang, waydtgang, wäydtgang, weidganng, weydgang, weydtgang *m Untergang*, *Grenzbegehung* 286²⁶⁻³⁹, 500^{3,8}, 1089³⁶, 1090¹⁵⁻³⁸, 1092⁷

weidgang weydgang, weydtganng *m Weidgang*, *Weiderecht* 617¹⁶, 656³⁶, 657⁴¹, 1007³³, 1008²³

Weidgang, *Grenzbegehung* 105⁴⁵, 207¹⁴, 435³⁵, 822³, 1089^{4,28}, 1090²⁹, 1091⁷,

Weihe 145²⁹, 412⁹, *s. auch* kirchweichungen, kirchwichungßtag, weichen

Weiher 19⁵, 182⁴, *s. auch* weyerstatt, wyer

wein win, wyn, wynn *m Wein* 7³³, 9²⁴, 23³, 25⁶, 82⁹, 163³⁹, 166⁴¹, 170¹¹, 183²⁹, 189²⁶, 260²¹, 261³⁸⁻⁴¹, 347⁵⁻¹¹, 349^{10,28}, 350³⁹, 353³⁴, 389²⁷, 397³⁵, 421²⁸, 425³⁰, 428³⁰, 499³⁶, 501⁵, 504²¹, 553^{14,26}, 590^{2,15}, 605⁴, 614^{28,34}, 630⁷, 677³⁹, 684^{30,31}, 685³³, 690¹⁵, 705^{6,11}, 726^{12,13}, 734²⁷, 739^{25,31}, 740¹⁵, 743¹⁶, 749³⁶, 760⁴⁴, 761²⁶, 762³⁴, 771⁹⁻³⁶, 773⁶⁻⁴⁵, 774⁹⁻³⁷, 775^{6,7}, 794^{22,29}, 797³⁸, 799⁷, 805⁶, 832^{6,7}, 833⁴, 843^{45,47}, 844^{18,19}, 849¹¹, 885²⁸, 888¹⁶, 913³, 957¹³, 968³², 994¹⁰, 1015⁷, 1032¹⁷, 1078^{7,10}, 1163¹⁹

Wein 24²⁰, 313¹², 375⁹, 396²¹, 399¹⁷, 646²⁵, 688³², 740¹⁸, 763², 771^{1,31}, 775¹⁹, 798^{30,40}, 833^{14,29}, 844²⁸⁻⁴⁸, 845^{17,24}, 957¹⁶, 958¹, 1063⁴⁰, 1074⁴⁸, 1078^{3,24}, 1217⁶, *s. auch Branntwein*, landwein, schenkwin, *Torkel*, Überreiner, *Ungeld (Umgeld)*, Veltliner, wein greüweli, wein, weingartman, weingeschir, weinkauf, weinmost, weinschäzer, welt-schen win, winfürer, winher

wein greüweli *m schlechter Wein* 1078²⁴

Verortung des SSRQ-Vokabulars

Seminararbeit von Dr. [Cathrin Senn](#)

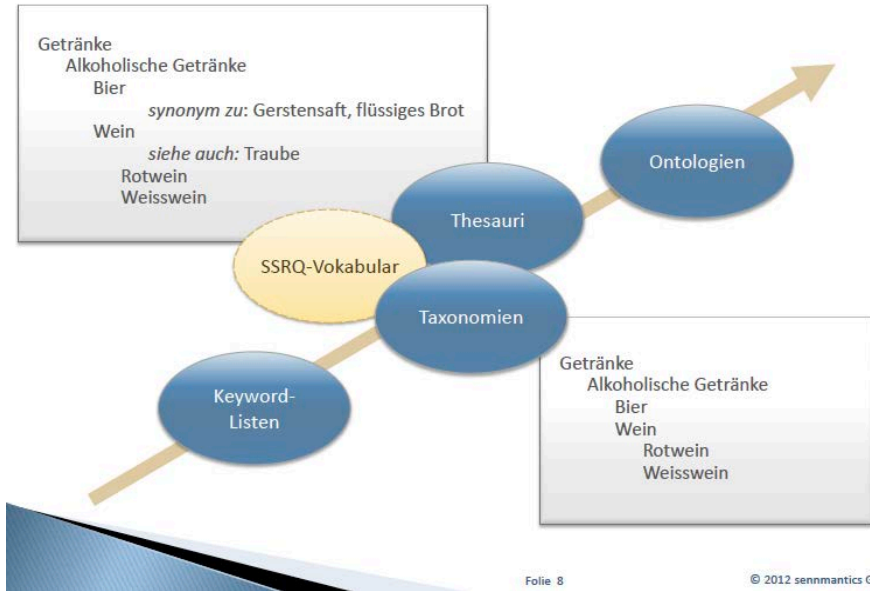
- 22 jüngsten SSRQ-Bände
- ZH, BE, LU, FR, AI/AR, SG, GR, AG, VD
- Zeitraum von 1100 Jahren
- 70 000 Einträge, 5100 deutsche Schlagworte

Verortung des SSRQ-Vokabulars

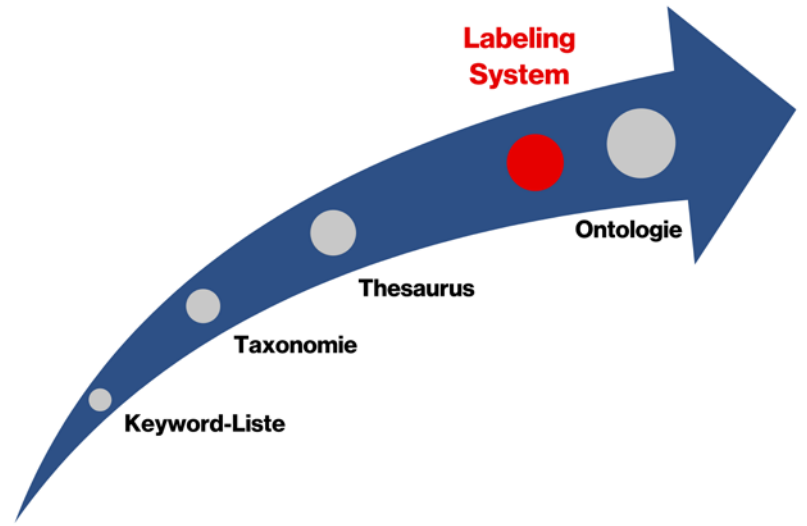
- 15% rechtlich relevante Schlagworte
- viele Berufs-, Amts- und Rollenbezeichnungen (Funktion)
- 9% auch in HRG



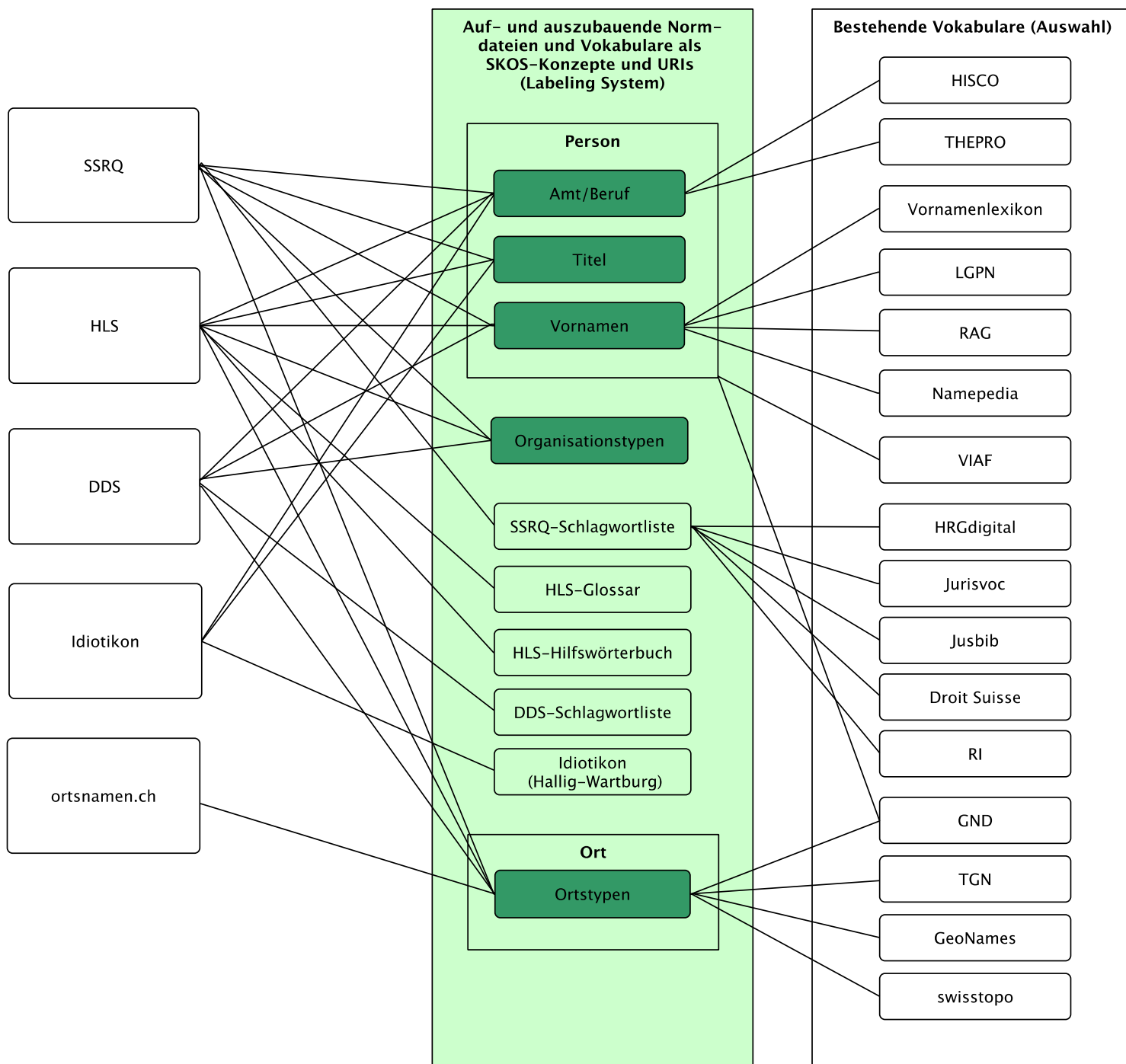
Verortung Vokabulare



Graphiken: Cathrin Senn, Florian Thiery



zwischen Thesaurus und Ontologie.
«enriched thesaurus»



Label

Ressource / String

mandatory

optional

<http://ssrq-sds-fds.ch/online/LU_II_3>
<http://ssrq-sds-fds.ch/online/SG_III_2>

normalisierter Vorname

normalisierter Vorname

name@lang (Synonym)

name@lang

Onomastik

DBpedia

forename@lang-std/orig

forename@lang-std/orig

...

Vornamenvarianten

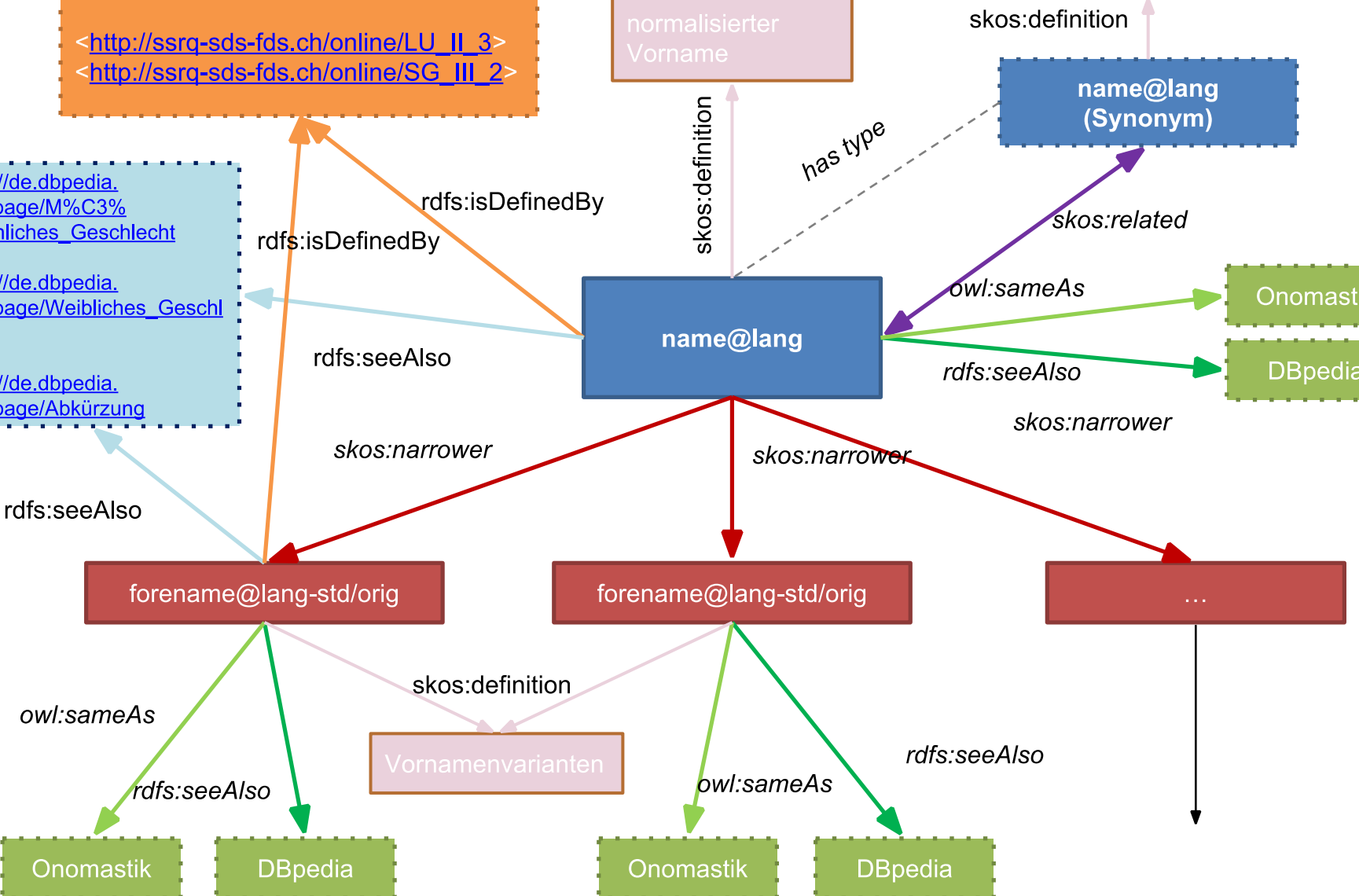
Onomastik

DBpedia

Onomastik

DBpedia

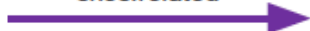
http://de.dbpedia.org/page/M%C3%A4nnliches_Geschlecht
http://de.dbpedia.org/page/Weibliches_Geschlecht
<http://de.dbpedia.org/page/Abkürzung>



Legende

Beispiel: Johann (Auszug)

skos:related



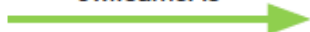
skos:narrower



rdfs:isDefinedBy



owl:sameAs



rdfs:seeAlso



Label

Ressource /
String

etc.

Johann@de

Giovanni@it

http://dbpedia.org/page/Männliches_Geschlecht

<http://dbpedia.org/page/Hans>

Hans@de

Johannes@de

<http://dbpedia.org/page/Johannes>

http://www.onomastik.com/Vornamen-Lexikon/name_65_Hans.html

http://www.onomastik.com/Vornamen-Lexikon/name_86_Johannes.html

http://ssrq-sds-fds.ch/online/SG_III_2

Hans@de-orig

http://ssrq-sds-fds.ch/online/SG_III_2

Haniß@de-orig

http://ssrq-sds-fds.ch/online/SG_III_2

Hanns@de-orig

http://ssrq-sds-fds.ch/online/LU_II_3

Hannß@de-orig

http://ssrq-sds-fds.ch/online/SG_III_2

Hanss@de-orig

http://ssrq-sds-fds.ch/online/LU_II_3

Hanß@de-orig

http://ssrq-sds-fds.ch/online/SG_III_2

http://dbpedia.org/page/Männliches_Geschlecht

Johannes@de-orig

Johanneß@de-orig

Johanniß@de-orig

Johanns@de-orig

Johnnis@de-orig

http://ssrq-sds-fds.ch/online/SG_III_2

http://ssrq-sds-fds.ch/online/SG_III_2

http://ssrq-sds-fds.ch/online/SG_III_2

http://ssrq-sds-fds.ch/online/SG_III_2

http://ssrq-sds-fds.ch/online/SG_III_2

http://ssrq-sds-fds.ch/online/LU_II_3

http://ssrq-sds-fds.ch/online/LU_II_3

Vornamen-Suche: Demo «Johann»

Durch Nutzung des LS SPARQL Endpoints

<http://embed.plnkr.co/zGjOuT/preview>



Ergebnisse der Probephase

- Vornamen der SSRQ als Ontologie modelliert
- Workflow der Datenaufbereitung entwickelt
- Erfolgreicher Upload ins LS
- Abfrageschnittstelle entwickelt zur Illustration der Nutzungsmöglichkeiten des LS
- Überprüfung der Vornamen-Normdatei dank Abfrage
- Erfahrungen gesammelt mit dem gut konzipierten Tool LS



Danke – Fragen?

Applikation LS-Prototyp:

<http://labeling.i3mainz.hs-mainz.de/>

