An Evaluation of RDF Constraint Types to Validate Metadata on Person-Level and Aggregated Data

Thomas Bosch[1], Benjamin Zapilko[1], Joachim Wackerow[1], and Kai Eckert[2]

[1] GESIS – Leibniz Institute for the Social Sciences, Germany
{firstname.lastname}@gesis.org,
[2] University of Mannheim, Germany
kai@informatik.uni-mannheim.de

Abstract. ...

Keywords: RDF Validation, RDF Constraints, DDI-RDF Discovery Vocabulary, Disco, RDF Data Cube Vocabulary, Linked Data, Semantic Web

1 Evaluation

	Vocabularies				
Criteria (Counts)	**Person-Level Metadata (Disco)**	**Aggregated Metadata (Data Cube)**	**Thesauri (SKOS)**	**Rectangular Data (PHDD)**	**Statistical Classifications (XKOS)** **Total**
Validated Triples	9,673,055	3,775,983,610	477,737,281		
Validated Data Sets	1,526	9,990	4,178		
Constraint Violations	XXXXX				
Constraint Violations ($\mathcal{SL}_0$)	XXXXX				
Constraint Violations ($\mathcal{SL}_1$)	473,574 ≡ XXXXX%				
Constraint Violations ($\mathcal{SL}_2$)	XXXXX				
Constraint Types	52 (15|37)				
Constraint Types ($\mathcal{C}_C$)	30 ≡ 57.7 %				
Constraint Types ($\mathcal{C}_T$)	22 ≡ 42.3 %				
Constraints	142 (77|65)				
Constraints ($\mathcal{C}_C$)	72 ≡ 50.7 %				
Constraints ($\mathcal{C}_T$)	70 ≡ 49.3 %				
Constraints ($\mathcal{SL}_0$)	75 ≡ 52.8% (44|31)				
Constraints ($\mathcal{SL}_1$)	9 ≡ 6.3% (8|1)				
Constraints ($\mathcal{SL}_2$)	58 ≡ 40.8% (25|33)				

Table 1: Evaluation

1.1 Legend

Validation Successful: X indicates the number of raised constraint violation triples.

Symbol	Description
X	Validation Successful
>X	Poor Performance/Scaling
✗	Very Poor Performance/Scaling
(!)	Not Yet Implemented Constraint
(X)	The validation of X data sets could not be finished due to SPARQL endpoints' technical restrictions (e.g. the defined timeout).
*	severity level $\mathcal{SL}_0$ (informational)
**	severity level $\mathcal{SL}_1$ (warning)
***	severity level $\mathcal{SL}_2$ (error)

Table 2: Legend

Poor Performance/Scaling: The performance of the implementation of the underlying SPARQL CONSTRUCT query is too poor to get all resulting constraint violation triples. Therefore, a limit of X result constraint violation triples is set. It is likely that there are more than X constraint violations. Although, the result set contains not the whole set of raised constraint violation triples, the constraint can be used as an indicator if there data not conforming to the constraint and to resolve constraint violations step by step. As part of future work, the performance will be improved.

Very Poor Performance/Scaling: The performance of the implementation of the underlying SPARQL CONSTRUCT query is too poor to get any results, even though a limit of result constraint violation triples is set. As part of future work, the performance will be improved.

2 Person-Level Metadata (Disco)

2.1 Evaluation Results

Legend: total (implemented|not yet implemented)

Evaluation Criteria	**Counts**
Validated Triples	9,673,055
Validated Data Sets	1,526
Constraint Violations	XXXXX
Constraint Violations ($\mathcal{SL}_0$)	XXXXX
Constraint Violations ($\mathcal{SL}_1$)	473,574 ≡ XXXXX%
Constraint Violations ($\mathcal{SL}_2$)	XXXXX
Constraint (Most Constraint Violations)	*DISCO-C-LABELING-AND-DOCUMENTATION-06* (547,916)
	DISCO-C-COMPARISON-VARIABLES-02 (547,916)
Constraint (Most Constraint Violations ($\mathcal{SL}_0$))	*DISCO-C-LABELING-AND-DOCUMENTATION-06* (547,916)
Constraint (Most Constraint Violations ($\mathcal{SL}_1$))	*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-46* (468,807)
Constraint (Most Constraint Violations ($\mathcal{SL}_2$))	*DISCO-C-COMPARISON-VARIABLES-02* (547,916)
Constraint Types	52 (15|37)
Constraint Types ($\mathcal{C}_C$)	30 ≡ 57.7 %
Constraint Types ($\mathcal{C}_T$)	22 ≡ 42.3 %
Constraint Types (Most Constraints)	1. Existential Quantifications: 46 ≡ 32.4% (46|0)
	2. Data Model Consistency: 7 (1|6)
	3. Aggregation: 7 (0|7)
Constraint Type (Most Constraints ($\mathcal{SL}_2$))	Existential Quantifications: 9 (9|0)
Constraints	142 (77|65)
Constraints ($\mathcal{C}_C$)	72 ≡ 50.7 %
Constraints ($\mathcal{C}_T$)	70 ≡ 49.3 %
Constraints ($\mathcal{SL}_0$)	75 ≡ 52.8% (44|31)
Constraints ($\mathcal{SL}_1$)	9 ≡ 6.3% (8|1)
Constraints ($\mathcal{SL}_2$)	58 ≡ 40.8% (25|33)

Table 3: Evaluation of Disco Data Sets - Evaluation Results

2.2 Data Sets Overview

Abbr.	**Disco Data Sets**
Missy	*Microdata Information System*[3]
DwB	*DwB Discovery Portal*[4]
DDA-SND	*DDI-RDF*[5]
	provided by the *Danish Data Archive (DDA)*[6] and Swedish National Data Service (SND)[7]

Table 4: Disco Data Sets Abbreviations

	Counts									
Data Sets	**triples**	**disco:StudyGroup**	**disco:Study**	**disco:LogicalDataSet**	**disco:Universe**	**disco:Variable**	**disco:Question**	**disco:SummaryStatistics**	**disco:CategoryStatistics**	**skos:Concept**
Missy	5,068,838	6	45	159	1,125	21,040	0	0	0	147,193
DwB	2,332,802	0	1,387	1,367	2,796	446,806	0	0	0	0
DDA-SND	2,271,415	0	1,490	0	10,188	80,070	139,237	0	0	290,963
Total	9,673,055			1,526						

Table 5: Disco Data Sets Overview

Data Sets	**SPARQL Endpoint**
Missy	http://svko-missy:8181/openrdf-workbench/repositories/native-java-store/summary
DwB	http://dwb-dev.nsd.uib.no/sparql
DDA-SND	http://ddi-rdf.borsna.se/endpoint/

Table 6: Disco SPARQL Endpoints

[3] http://www.gesis.org/missy/eu/missy-home
[4] http://dwb-dev.nsd.uib.no/portal
[5] http://ddi-rdf.borsna.se/
[6] http://samfund.dda.dk/dda/default-en.asp
[7] http://snd.gu.se/en

2.3 Detailed Evaluation

	Data Sets		
Existential Quantifications (1)	*Missy*	*DwB*	*DDA-SND*
DISCO-C-EXISTENTIAL-QUANTIFICATIONS-01***	✓	✓	✓
DISCO-C-EXISTENTIAL-QUANTIFICATIONS-02***	7	17	1,490
DISCO-C-EXISTENTIAL-QUANTIFICATIONS-03*	✓	✓	✓
DISCO-C-EXISTENTIAL-QUANTIFICATIONS-04*	11,021	445,381	62,260
DISCO-C-EXISTENTIAL-QUANTIFICATIONS-05*	✓	✓	139,237
DISCO-C-EXISTENTIAL-QUANTIFICATIONS-06*	12	1,367	✓
DISCO-C-EXISTENTIAL-QUANTIFICATIONS-07*	6	✓	✓
DISCO-C-EXISTENTIAL-QUANTIFICATIONS-08*	45	1,387	1,490
DISCO-C-EXISTENTIAL-QUANTIFICATIONS-09*	6	✓	✓
DISCO-C-EXISTENTIAL-QUANTIFICATIONS-10*	45	1,387	1,490

Table 7: Evaluation of Disco Data Sets - Existential Quantifications (1)

	Data Sets		
Existential Quantifications (2)	*Missy*	*DwB*	*DDA-SND*
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-11**	6	✓	✓
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-12**	6	✓	✓
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-13**	✓	✓	✓
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-14**	45	1,387	1,490
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-15**	45	1,387	1,490
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-16**	✓	✓	✓
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-17**	159	1,367	✓
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-18**	159	1,367	✓
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-19**	✓	✓	✓
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-20**	✓	1,367	✓

Table 8: Evaluation of Disco Data Sets - Existential Quantifications (2)

	Data Sets		
Existential Quantifications (3)	*Missy*	*DwB*	*DDA-SND*
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-21**	✓	1,367	✓
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-22**	✓	✓	✓
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-23**	6	✓	✓
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-24**	45	1,387	1,490
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-25**	45	1,387	1,490
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-26**	45	1,387	1,490
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-27****	✓	130	1,490
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-28***	159	✓	✓
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-29***	✓	✓	✓
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-30***	✓	✓	✓

Table 9: Evaluation of Disco Data Sets - Existential Quantifications (3)

Existential Quantifications (4)	Data Sets		
	Missy	*DwB*	*DDA-SND*
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-31***	159	1,367	✓
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-32****	✓	✓	✓
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-33****	✓	✓	✓
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-34****	✓	✓	✓
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-35****	✓	✓	✓
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-36****	✓	✓	✓
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-37**	18,625	✓	✓
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-38**	✓	✓	750
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-39****	✓	✓	✓
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-40**	✓	✓	139,237

Table 10: Evaluation of Disco Data Sets - Existential Quantifications (4)

Existential Quantifications (5)	Data Sets		
	Missy	*DwB*	*DDA-SND*
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-41**	✓	✓	✓
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-42**	✓	✓	✓
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-43**	15,733	446,806	80,070
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-44**	159	✓	✓
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-45**	6,784	446,806	19,221
*DISCO-C-EXISTENTIAL-QUANTIFICATIONS-46***	11,550	446,806	10,451

Table 11: Evaluation of Disco Data Sets - Existential Quantifications (5)

	Data Sets		
Conditional Properties	*Missy*	*DwB*	*DDA-SND*
DISCO-C-CONDITIONAL-PROPERTIES-01***	✓	✓	80,070
DISCO-C-CONDITIONAL-PROPERTIES-02**	12	✓	✓
DISCO-C-CONDITIONAL-PROPERTIES-03**	90	✓	2,980
DISCO-C-CONDITIONAL-PROPERTIES-04***	6	✓	✓
DISCO-C-CONDITIONAL-PROPERTIES-05***	45	1,387	1,490
DISCO-C-CONDITIONAL-PROPERTIES-06***	✓	✓	✓

Table 12: Evaluation of Disco Data Sets - Conditional Properties

	Data Sets		
Provenance	*Missy*	*DwB*	*DDA-SND*
DISCO-C-PROVENANCE-01*	6	✓	✓
DISCO-C-PROVENANCE-02*	45	1,387	1,490
DISCO-C-PROVENANCE-03*	159	1,367	✓
DISCO-C-PROVENANCE-04*	✓	1,367	✓

Table 13: Evaluation of Disco Data Sets - Provenance

	Data Sets		
Labeling and Documentation	*Missy*	*DwB*	*DDA-SND*
DISCO-C-LABELING-AND-DOCUMENTATION-01[*]	6	✓	✓
DISCO-C-LABELING-AND-DOCUMENTATION-02[*]	45	1,387	1,490
DISCO-C-LABELING-AND-DOCUMENTATION-03[*]	159	1,367	✓
DISCO-C-LABELING-AND-DOCUMENTATION-04[*]	✓	1,367	✓
DISCO-C-LABELING-AND-DOCUMENTATION-05[*]	✓	✓	✓
DISCO-C-LABELING-AND-DOCUMENTATION-06[*]	21,040	446,806	80,070

Table 14: Evaluation of Disco Data Sets - Labeling and Documentation

	Data Sets		
Data Model Consistency	*Missy*	*DwB*	*DDA-SND*
DISCO-C-DATA-MODEL-CONSISTENCY-01 (!)[***]			
DISCO-C-DATA-MODEL-CONSISTENCY-02 (!)[***]			
DISCO-C-DATA-MODEL-CONSISTENCY-03 (!)[***]			
DISCO-C-DATA-MODEL-CONSISTENCY-04 (!)[***]			
DISCO-C-DATA-MODEL-CONSISTENCY-05[***]	✓	✓	✓
DISCO-C-DATA-MODEL-CONSISTENCY-06 (!)[***]			
DISCO-C-DATA-MODEL-CONSISTENCY-07 (!)[***]			

Table 15: Evaluation of Disco Data Sets - Data Model Consistency

	Data Sets		
Comparison	*Missy*	*DwB*	*DDA-SND*
*DISCO-C-COMPARISON-VARIABLES-01 (!)***			
*DISCO-C-COMPARISON-VARIABLES-02****	21,040	446,806	80,070
*DISCO-C-COMPARISON-VARIABLES-03 (!)****			
*DISCO-C-COMPARISON-VARIABLES-04**	18,625	✓	✓
*DISCO-C-COMPARISON-VARIABLES-05****	159	✓	✓

Table 16: Evaluation of Disco Data Sets - Comparison

	Data Sets		
Mathematical Operations	*Missy* *DwB*		*DDA-SND*
*DISCO-C-MATHEMATICAL-OPERATIONS-01 (!)****			
*DISCO-C-MATHEMATICAL-OPERATIONS-02 (!)****			
*DISCO-C-MATHEMATICAL-OPERATIONS-03 (!)****			
*DISCO-C-MATHEMATICAL-OPERATIONS-04 (!)****			
*DISCO-C-MATHEMATICAL-OPERATIONS-05 (!)****			

Table 17: Evaluation of Disco Data Sets - Mathematical Operations

	Data Sets	
Language Tags	*Missy DwB*	*DDA-SND*
*DISCO-C-LANGUAGE-TAG-MATCHING-01 (!)**		
*DISCO-C-LANGUAGE-TAG-CARDINALITY-01 (!)**		
*DISCO-C-LANGUAGE-TAG-CARDINALITY-02 (!)**		
*DISCO-C-LANGUAGE-TAG-CARDINALITY-03 (!)**		

Table 18: Evaluation of Disco Data Sets - Language Tags

	Data Sets	
Aggregation	*Missy DwB*	*DDA-SND*
*DISCO-C-AGGREGATION-01 (!)**		
*DISCO-C-AGGREGATION-02 (!)**		
*DISCO-C-AGGREGATION-03 (!)**		
*DISCO-C-AGGREGATION-04 (!)**		
*DISCO-C-AGGREGATION-05 (!)**		
*DISCO-C-AGGREGATION-06 (!)**		
*DISCO-C-AGGREGATION-07 (!)**		

Table 19: Evaluation of Disco Data Sets - Aggregation

	Data Sets		
Disco Constraints	*Missy*	*DwB*	*DDA-SND*
DISCO-C-ALLOWED-VALUES-01***	✓	✓	✓
DISCO-C-LITERAL-RANGES-01***	✓	✓	✓
DISCO-C-INVERSE-FUNCTIONAL-PROPERTIES-01***	✓	✓	✓
DISCO-C-INVERSE-FUNCTIONAL-PROPERTIES-02***	✓	✓	✓
DISCO-C-CLASS-SPECIFIC-PROPERTY-RANGE-01***	✓	✓	✓
DISCO-C-MEMBERSHIP-IN-CONTROLLED-VOCABULARIES-01***	✓	✓	✗
DISCO-C-LITERAL-VALUE-COMPARISON-01***	✓	1,299	✓
DISCO-C-CONTEXT-SPECIFIC-VALID-PROPERTIES-01*	21,038	✓	✓
DISCO-C-DATA-PROPERTY-FACETS-01**	✓	✓	✓
DISCO-C-DATA-PROPERTY-FACETS-02**	✓	✓	✓

Table 20: Evaluation of Disco Data Sets - Disco Constraints (1)

	Data Sets		
Disco Constraints	*Missy*	*DwB*	*DDA-SND*
DISCO-C-VALUE-IS-VALID-FOR-DATATYPE-01***	30	6,932	✓
DISCO-C-VALUE-IS-VALID-FOR-DATATYPE-02***	✓	✓	✓
DISCO-C-SUBSUMPTION-01 (!)***			
DISCO-C-CLASS-EQUIVALENCE-01 (!)*			
DISCO-C-SUB-PROPERTIES-01 (!)***			
DISCO-C-PROPERTY-DOMAIN-01 (!)***			
DISCO-C-PROPERTY-RANGES-01 (!)***			
DISCO-C-INVERSE-OBJECT-PROPERTIES-01 (!)***			
DISCO-C-INVERSE-OBJECT-PROPERTIES-02 (!)***			
DISCO-C-INVERSE-OBJECT-PROPERTIES-03 (!)***			
DISCO-C-DISJOINT-PROPERTIES-01 (!)***			

Table 21: Evaluation of Disco Data Sets - Disco Constraints (2)

	Data Sets	
Disco Constraints	*Missy DwB*	*DDA-SND*
DISCO-C-ASYMMETRIC-OBJECT-PROPERTIES-01 (!)***		
DISCO-C-IRREFLEXIVE-OBJECT-PROPERTIES-01 (!)***		
DISCO-C-CLASS-SPECIFIC-IRREFLEXIVE-OBJECT-PROPERTIES-01 (!)***		
DISCO-C-CLASS-SPECIFIC-IRREFLEXIVE-OBJECT-PROPERTIES-02 (!)***		
DISCO-C-DISJOINT-CLASSES-01 (!)***		
DISCO-C-EQUIVALENT-PROPERTIES-01 (!)*		
DISCO-C-LITERAL-PATTERN-MATCHING-01 (!)*		
DISCO-C-DISJUNCTION-01 (!)***		
DISCO-C-UNIVERSAL-QUANTIFICATIONS-01 (!)***		
DISCO-C-MINIMUM-QUALIFIED-CARDINALITY-RESTRICTIONS-01 (!)***		

Table 22: Evaluation of Disco Data Sets - Disco Constraints (3)

Disco Constraints	Data Sets: *Missy DwB*	Data Sets: *DDA-SND*
DISCO-C-MAXIMUM-QUALIFIED-CARDINALITY-RESTRICTIONS-01 (!)***		
DISCO-C-EXACT-QUALIFIED-CARDINALITY-RESTRICTIONS-01 (!)***		
DISCO-C-CONTEXT-SPECIFIC-EXCLUSIVE-OR-OF-PROPERTY-GROUPS-01 (!)*		
DISCO-C-IRI-PATTERN-MATCHING-01 (!)*		
DISCO-C-ORDERING-01 (!)*		
DISCO-C-ORDERING-02 (!)*		
DISCO-C-ORDERING-03 (!)*		
DISCO-C-STRING-OPERATIONS-01 (!)*		
DISCO-C-CONTEXT-SPECIFIC-VALID-CLASSES-01 (!)*		
DISCO-C-CONTEXT-SPECIFIC-VALID-PROPERTIES-01 (!)*		

Table 23: Evaluation of Disco Data Sets - Disco Constraints (4)

	Data Sets	
Disco Constraints	*Missy DwB*	*DDA-SND*
DISCO-C-DEFAULT-VALUES-01 (!)[*]		
DISCO-C-WHITESPACE-HANDLING-01 (!)[*]		
DISCO-C-HTML-HANDLING-01 (!)[*]		
DISCO-C-HTML-HANDLING-02 (!)[*]		
DISCO-C-RECOMMENDED-PROPERTIES-01 (!)[*]		
DISCO-C-HANDLE-RDF-COLLECTIONS-01 (!)[*]		
DISCO-C-HANDLE-RDF-COLLECTIONS-02 (!)[*]		
DISCO-C-USE-SUB-SUPER-RELATIONS-IN-VALIDATION-01 (!)[*]		
DISCO-C-USE-SUB-SUPER-RELATIONS-IN-VALIDATION-02 (!)[*]		
DISCO-C-STRUCTURE-01 (!)[***]		

Table 24: Evaluation of Disco Data Sets - Disco Constraints (5)

	Data Sets	
Disco Constraints	*Missy DwB*	*DDA-SND*
DISCO-C-VOCABULARY-01 (!)[***]		
DISCO-C-HTTP-URI-SCHEME-VIOLATION (!)[***]		

Table 25: Evaluation of Disco Data Sets - Disco Constraints (6)

3 Aggregated Metadata (Data Cube)

3.1 Evaluation Results

Evaluation Criteria	**Counts**
Validated Triples	3,775,983,610
Validated Data Sets	9,990
Constraint Violations	45,635,846
Constraint Violations ($\mathcal{SL}_0$)	0 ≡ 0%
Constraint Violations ($\mathcal{SL}_1$)	45,520,613 ≡ 99.75%
Constraint Violations ($\mathcal{SL}_2$)	115,233 ≡ 0.25%
Constraint (Most Constraint Violations)	*DATA-MODEL-CONSISTENCY-05* (45,514,102)
Constraint (Most Constraint Violations ($\mathcal{SL}_0$))	-
Constraint (Most Constraint Violations ($\mathcal{SL}_1$))	*DATA-MODEL-CONSISTENCY-05* (45,514,102)
Constraint (Most Constraint Violations ($\mathcal{SL}_2$))	*MINIMUM-QUALIFIED-CARDINALITY-RESTRICTIONS-02* (1,556)
Constraint Types	20 (7|13)
Constraint Types ($\mathcal{C}_C$)	5 ≡ 25 %
Constraint Types ($\mathcal{C}_T$)	15 ≡ 75 %
Constraint Types (Most Constraints)	1. Data Model Consistency: 11 ≡ 31.4% (10|1) 2. Existential Quantifications: 4 ≡ 11.4% (4|0)
Constraint Type (Most Constraints ($\mathcal{SL}_2$))	Data Model Consistency: 8 ≡ 22.9% (7|1)
Constraints	35 (20|15)
Constraints ($\mathcal{C}_C$)	16 (12|4)
Constraints ($\mathcal{C}_T$)	19 (8|11)
Constraints ($\mathcal{SL}_0$)	4 ≡ 11.4% (0|4)
Constraints ($\mathcal{SL}_1$)	3 ≡ 8.6% (3|0)
Constraints ($\mathcal{SL}_2$)	28 ≡ 80% (17|11)

Table 26: Evaluation of Data Cube Data Sets - Evaluation Results

3.2 Data Sets Overview

- overview over Data Cube data sets[8]
- http://ontologycentral.com/

[8] http://270a.info/; http://datahub.io/de/dataset?tags=format-qb

Abbr.	Data Cube Data Sets
ECB	*European Central Bank*[9]
UIS	*UNESCO Institute for Statistics*[10]
IMF	*International Monetary Fund*[11]
BFS	*Bundesamt für Statistik - Swiss Federal Statistics*[12]
FAO	*Food and Agriculture Organization of the United Nations*[13]
WB	*World Bank*[14]
FRB	*Federal Reserve Board*[15]
TI	*Transparency International*[16]
OECD	*Organisation for Economic Co-operation and Development*[17]
BIS	*Bank for International Settlements*[18]
ABS	*Australian Bureau of Statistics*[19]
IEEE-VIS	*IEEE VIS Source Data*
ACORN-SAT	*Australian Climate Observations Reference Network - Surface Air Temperature Dataset*
HDP	*HealthData.gov Platform (HDP) on the Semantic Web*
Eurostat	*The Eurostat Linked Data* (SPARQL endpoint unavailable)
Asturias	*Nomenclator Asturias* (SPARQL endpoint unavailable!)
ISTAT	*ISTAT Immigration (LinkedOpenData.it)* (SPARQL endpoint unavailable)
ICANE	*Statistical Office of Cantabria (Instituto Cántabro de Estadística, ICANE)* (SPARQL endpoint unavailable)
EE-2009	*European Election Results 2009* (SPARQL endpoint unavailable)
EU-B	*Standard Eurobarometer* (SPARQL endpoint unavailable)
ECB-S	*European Central Bank Statistics (PublicData.eu)* (SPARQL endpoint unavailable)
CPV-2008	*Common Procurement Vocabulary (CPV) 2008* (SPARQL endpoint unavailable)
CPV-2003	*Common Procurement Vocabulary (CPV) 2003* (SPARQL endpoint unavailable)

Table 27: Data Cube Data Sets Abbreviations

[9] http://www.ecb.europa.eu/home/html/index.en.html
[10] http://www.uis.unesco.org/Pages/default.aspx
[11] http://www.imf.org/external/index.htm
[12] http://www.bfs.admin.ch/
[13] http://www.fao.org/home/en/
[14] http://www.worldbank.org/
[15] http://www.federalreserve.gov/
[16] http://www.transparency.org/
[17] http://www.oecd.org/
[18] http://www.bis.org/
[19] http://abs.gov.au/

	Counts				
Data Sets	**triples**	**qb:DataSet**	**qb:DataStructureDefinition**	**qb:Observation**	**qb:Slice**
ECB	468,899,474	55	46	>11,000,000	428,698
UIS	10,400,534	5	5	1,437,651	0
IMF	35,688,446	4	8	3,603,719	0
BFS	1,533,743	0	0	8	0
FAO	53,000,000	10	10	>7,100,000	0
WB	174,006,552	9,466	59	>17,000,000	0
FRB	185,266,900	49	98	>9,500,000	0
TI	52,233	6	6	3,928	0
OECD	304,995,160	136	140	>12,000,000	0
BIS	54,197,482	6	12	3,606,466	47,914
ABS	2,357,400,000	253	257	>11,000,000	0
IEEE-VIS	19,935,340	0	0	1,350	0
ACORN-SAT	98,381,319	0	4	0	0
HDP	12,226,427	0	0	0	0
Total	3,775,983,610	9,990			

Table 28: Data Cube Data Sets Overview

Data Sets	SPARQL Endpoints
ECB	http://ecb.270a.info/sparql
UIS	http://uis.270a.info/sparql
IMF	http://imf.270a.info/sparql
BFS	http://bfs.270a.info/sparql
FAO	http://fao.270a.info/sparql
WB	http://worldbank.270a.info/sparql
FRB	http://frb.270a.info/sparql
TI	http://transparency.270a.info/sparql
OECD	http://oecd.270a.info/sparql
BIS	http://bis.270a.info/sparql
ABS	http://abs.270a.info/sparql
ACORN-SAT	http://lab.environment.data.gov.au/sparql
HDP	http://healthdata.tw.rpi.edu/sparql

Table 29: Data Cube SPARQL Endpoints

3.3 Detailed Evaluation

	Data Sets						
Data Model Consistency	*ECB*	*UIS*	*IMF*	*BFS*	*FAO*	*WB*	*FRB*
*DATA-MODEL-CONSISTENCY-01***	✓ (2)	✓	✓	✓	✓	✓	✓
*DATA-MODEL-CONSISTENCY-02****	✓ (2)	✓	✓	✓	✓	✓	✓
*DATA-MODEL-CONSISTENCY-03****	✓ (2)	✓	✓	✓	✓	✓	✓
*DATA-MODEL-CONSISTENCY-04****	✓ (6)	✓	✓	✓	✓	✓	14,372
*DATA-MODEL-CONSISTENCY-05***	1,198,352 (50)	✗	✗	✓	✗	✓	16,175,814 (42)
*DATA-MODEL-CONSISTENCY-06****	✓ (2)	✓	✓	✓	✓	✓	✓
*DATA-MODEL-CONSISTENCY-07****	✓ (9)	✓	99,091	✓	✓	✓	✓ (1)
*DATA-MODEL-CONSISTENCY-08****	✓ (2)	✓	✓	✓	✓	✓	✓
*DATA-MODEL-CONSISTENCY-09****	✓ (2)	✓	✓	✓	✓	✓	✓
*DATA-MODEL-CONSISTENCY-10**** (!)	-	-	-	-	-	-	-
*DATA-MODEL-CONSISTENCY-11***	6,511 (10)	✓	✓	✓	✓	✓	✓

Table 30: Evaluation of Data Cube Data Sets - Data Model Consistency (1)

	Data Sets						
Data Model Consistency	*TI*	*OECD*	*BIS*	*ABS*	*IEEE-VIS*	*ACORN-SAT*	*HDP*
*DATA-MODEL-CONSISTENCY-01***	✓	✓	✓	✓	✓	✓	✓
*DATA-MODEL-CONSISTENCY-02****	✓	✓	✓	✓	✓	8	✓
*DATA-MODEL-CONSISTENCY-03****	✓	✓	✓	✓	✓	✓	✓
*DATA-MODEL-CONSISTENCY-04****	✓	✓	✓	✓ (6)	✓	✓	✓
*DATA-MODEL-CONSISTENCY-05***	✓	21,142,838 (116)	✗	6,997,098 (246)	✓	✓	✓
*DATA-MODEL-CONSISTENCY-06****	✓	✓	✓	✓	✓	✓	✓
*DATA-MODEL-CONSISTENCY-07****	✓	✓	✓	✓ (8)	✓	✓	✓
*DATA-MODEL-CONSISTENCY-08****	✓	✓	✓	✓	✓	✓	✓
*DATA-MODEL-CONSISTENCY-09****	✓	✓	✓	✓	✓	✓	✓
*DATA-MODEL-CONSISTENCY-10**** (!)	-	-	-	-	-	-	-
*DATA-MODEL-CONSISTENCY-11***	✓	✓	✓	✓	✓	✓	✓

Table 31: Evaluation of Data Cube Data Sets - Data Model Consistency (2)

	Data Sets													
Existential Quantifications	*ECB*	*UIS*	*IMF*	*BFS*	*FAO*	*WB*	*FRB*	*TI*	*OECD*	*BIS*	*ABS*	*IEEE-VIS*	*ACORN-SAT*	*HDP*
*EXISTENTIAL-QUANTIFICATIONS-01****	9	✓	11	7	8	77	8	9	7	8	7	✓	✓	✓
*EXISTENTIAL-QUANTIFICATIONS-02****	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
*EXISTENTIAL-QUANTIFICATIONS-03****	✓	✓	✓	✓	✓	59	✓	6	✓	✓	✓	✓	4	✓
*EXISTENTIAL-QUANTIFICATIONS-04****	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Table 32: Evaluation of Data Cube Data Sets - Existential Quantifications

Cardinality Restrictions	Data Sets									
	ECB	*UIS*	*IMF*	*BFS*	*FAO*	*WB*	*FRB*	*TI*	*OECD*	*BIS*
MINIMUM-QUALIFIED-CARDINALITY-RESTRICTIONS-01 (!)[***]	-	-	-	-	-	-	-	-	-	-
MINIMUM-QUALIFIED-CARDINALITY-RESTRICTIONS-02[***]	✗	118	8	8	30	✓	30	✓	✗	12
MAXIMUM-QUALIFIED-CARDINALITY-RESTRICTIONS-01[***]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
EXACT-UNQUALIFIED-CARDINALITY-RESTRICTIONS-01[***]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
EXACT-QUALIFIED-CARDINALITY-RESTRICTIONS-02[***]	✓	✓	✓	✓	✓	1	✓	✓	✓	✓

Table 33: Evaluation of Data Cube Data Sets - Cardinality Restrictions (1)

Cardinality Restrictions	Data Sets			
	ABS	*IEEE-VIS*	*ACORN-SAT*	*HDP*
MINIMUM-QUALIFIED-CARDINALITY-RESTRICTIONS-01 (!)[***]	-	-	-	-
MINIMUM-QUALIFIED-CARDINALITY-RESTRICTIONS-02[***]	✗	1,350	✓	✓
MAXIMUM-QUALIFIED-CARDINALITY-RESTRICTIONS-01[***]	✓ (2)	✓	✓	✓
EXACT-UNQUALIFIED-CARDINALITY-RESTRICTIONS-01[***]	✓	✓	✓	✓
EXACT-QUALIFIED-CARDINALITY-RESTRICTIONS-02[***]	✓	✓	✓	✓

Table 34: Evaluation of Data Cube Data Sets - Cardinality Restrictions (2)

	Data Sets													
Structure	*ECB*	*UIS*	*IMF*	*BFS*	*FAO*	*WB*	*FRB*	*TI*	*OECD*	*BIS*	*ABS*	*IEEE-VIS*	*ACORN-SAT*	*HDP*
STRUCTURE-01***	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
STRUCTURE-02***	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Table 35: Evaluation of Data Cube Data Sets - Structure

	Data Sets													
Constraints	*ECB*	*UIS*	*IMF*	*BFS*	*FAO*	*WB*	*FRB*	*TI*	*OECD*	*BIS*	*ABS*	*IEEE-VIS*	*ACORN-SAT*	*HDP*
PROPERTY-DOMAIN-01 (!)***														
PROPERTY-RANGES-01 (!)***														
DISJOINT-PROPERTIES-01 (!)***														
DISJOINT-CLASSES-01 (!)***														
EQUIVALENT-PROPERTIES-01 (!)*														
UNIVERSAL-QUANTIFICATIONS-01 (!)***														
MEMBERSHIP-IN-CONTROLLED-VOCABULARIES-01 (!)***														
CONTEXT-SPECIFIC-VALID-CLASSES-01 (!)*														
CONTEXT-SPECIFIC-VALID-PROPERTIES-01 (!)*														
RECOMMENDED-PROPERTIES-01 (!)*														
VALUE-IS-VALID-FOR-DATATYPE-01 (!)***														
VOCABULARY-01 (!)***														

Table 36: Evaluation of Data Cube Data Sets - Constraints (1)

	Data Sets											
Constraints	*ECB*	*UIS*	*IMF*	*BFS*	*FAO*	*WB*	*FRB*	*TI*	*OECD*	*BIS*	*ABS*	*IEEE-VIS*	*ACORN-SAT*	*HDP*
HTTP-URI-SCHEME-VIOLATION (!)[***]														

Table 37: Evaluation of Data Cube Data Sets - Constraints (2)

4 Thesauri (SKOS)

4.1 Evaluation Results

Evaluation Criteria	**Counts**
Validated Triples	XXXXX
Validated Data Sets	XXXXX
Constraint Violations	XXXXX
Constraint Violations ($\mathcal{SL}_0$)	XXXXX
Constraint Violations ($\mathcal{SL}_1$)	XXXXX ≡ XXXXX%
Constraint Violations ($\mathcal{SL}_2$)	XXXXX
Constraint (Most Constraint Violations)	*LANGUAGE-TAG-CARDINALITY-01* (2,508,903)
Constraint (Most Constraint Violations ($\mathcal{SL}_0$))	*LABELING-AND-DOCUMENTATION-06* (1,022,362)
Constraint (Most Constraint Violations ($\mathcal{SL}_1$))	*LANGUAGE-TAG-CARDINALITY-01* (2,508,903)
Constraint (Most Constraint Violations ($\mathcal{SL}_2$))	-
Constraint Types	14 (4|10)
Constraint Types ($\mathcal{C}_C$)	5 ≡ 35.7%
Constraint Types ($\mathcal{C}_T$)	9 ≡ 64.3%
Constraint Types (Most Constraints)	1. Structure: 10 ≡ XX% (8|2)
	2. Labeling and Documentation: 6 ≡ XX% (5|1)
	3. Language Tag Cardinality: 4 ≡ XX% (4|0)
Constraint Type (Most Constraints ($\mathcal{SL}_2$))	Structure: 1 (0|1)
Constraints	35 (17|18)
Constraints ($\mathcal{C}_C$)	21 (13|8)
Constraints ($\mathcal{C}_T$)	14 (4|10)
Constraints ($\mathcal{SL}_0$)	21 ≡ 60% (12|9)
Constraints ($\mathcal{SL}_1$)	5 ≡ 14.3% (5|0)
Constraints ($\mathcal{SL}_2$)	9 ≡ 25.7% (0|9)

Table 38: Evaluation of Thesauri Data Sets - Evaluation Results

4.2 Data Sets Overview

- overview over SKOS data sets[20]
- overview over thesauri[21] (table 40)

[20] http://datahub.io/de/dataset?tags=format-skos
[21] http://datahub.io/de/dataset?tags=thesaurus

Abbr.	Thesauri
TheSoz	*Thesaurus for the Social Sciences*[22]
STW	*Thesaurus for Economics*[23]
AGROVOC	*AGROVOC Multilingual agricultural thesaurus*[24]
UNESCO	*UNESCO Thesaurus*[25]
TGN	*The Getty Thesaurus of Geographic Names*[26]
EARTh	*Environmental Applications Reference Thesaurus*[27]
ODT	*Open Data Thesaurus*[28]
SLD	*Spanish Linguistic Datasets*[29]
SSWT	*Social Semantic Web Thesaurus*[30]
GBA-GU	*Thesaurus of the Geological Survey of Austria (GBA) - Geology Unit*[31]
GBA-GTS	*Thesaurus of the Geological Survey of Austria (GBA) - Geologic Time Scale*[32]
GBA-L	*Thesaurus of the Geological Survey of Austria (GBA) - Lithology*[33]
GBA-LU	*Thesaurus of the Geological Survey of Austria (GBA) - Lithotectonic Unit*[34]
GEMET	*GEneral Multilingual Environmental Thesaurus*[35]
EuroVoc	*EuroVoc*[36]
CECCT	*Clean Energy and Climate Change Thesaurus*[37]

Table 39: Thesauri Abbreviations

[22] http://www.ecb.europa.eu/home/html/index.en.html
[23] http://zbw.eu/stw/versions/latest/about
[24] http://202.45.139.84:10035/catalogs/fao/repositories/agrovoc
[25] http://skos.um.es/sparql/
[26] http://vocab.getty.edu/sparql
[27] http://linkeddata.ge.imati.cnr.it/resource/EARTh/
[28] http://vocabulary.semantic-web.at/PoolParty/wiki/OpenData
[29] http://linguistic.linkeddata.es
[30] http://vocabulary.semantic-web.at/PoolParty/wiki/semweb
[31] http://resource.geolba.ac.at/
[32] http://resource.geolba.ac.at/
[33] http://resource.geolba.ac.at/
[34] http://resource.geolba.ac.at/
[35] http://www.eionet.europa.eu/gemet/
[36] http://open-data.europa.eu/de/data/dataset/eurovoc
[37] http://data.reegle.info/thesaurus/guide

	Counts						
Thesauri	**triples**	**skos:ConceptScheme**	**sko:Concept**	**skos:broader**	**skos:narrower**	**skos:hasTopConcept**	**skos:inScheme**
TheSoz	439,153	1	8,426	13,705	13,706	0	48,529
STW	221,668	1	13,468	13,732	13732	7	13,180
AGROVOC	6,080,477	1	32,310	33,507	33,507	25	32,310
UNESCO	288,346	9	26,714	20,028	20,028	607	32,009
TGN	16,112,321	8	2,898,775	0	0	0	1,453,767
EARTh	9,287,364	11	295,375	288,208	93,827	479	295,376
ODT	3,290	6	108	93	93	30	0
SLD	7,629,211	0	31,195	0	0	0	0
SSWT	64,698	9	2,127	2,300	2,301	38	0
GBA-GU	25,718	3	878	1,005	1,005	14	0
GBA-GTS	7,875	3	213	208	208	5	0
GBA-L	9,317	1	249	249	249	4	0
GBA-LU	9,504	3	364	359	359	7	0
GEMET	372,889,229	3,680	414,659	62,193	21,685	30,806	409,290
EuroVoc	64,477,774	439	79,557	6,922	0	532	14,428
CECCT	191,336	3	3,419	3,761	3,762	28	0
Total	477,737,281	4,178					

Table 40: Thesauri Overview

Thesauri	SPARQL Endpoints
TheSoz	http://lod.gesis.org/thesoz/sparql
STW	http://zbw.eu/beta/sparql/stw/query
AGROVOC	http://202.45.139.84:10035/catalogs/fao/repositories/agrovoc
UNESCO	http://skos.um.es/sparql/
TGN	http://vocab.getty.edu/
EARTh	http://linkeddata.ge.imati.cnr.it:8890/sparql
ODT	http://vocabulary.semantic-web.at/PoolParty/sparql/OpenData
SLD	http://linguistic.linkeddata.es/sparql
SSWT	http://vocabulary.semantic-web.at/PoolParty/sparql/semweb
GBA-GU	http://resource.geolba.ac.at/PoolParty/sparql/GeologicUnit
GBA-GTS	http://resource.geolba.ac.at/PoolParty/sparql/GeologicTimeScale
GBA-L	http://resource.geolba.ac.at/PoolParty/sparql/lithology
GBA-LU	http://resource.geolba.ac.at/PoolParty/sparql/tectonicunit
GEMET	http://semantic.eea.europa.eu/sparql
EuroVoc	http://open-data.europa.eu/de/linked-data
CECCT	http://poolparty.reegle.info/PoolParty/sparql/glossary

Table 41: Thesauri SPARQL Endpoints

4.3 Detailed Evaluation

	Data Sets									
Data Model Consistency	*TheSoz*	*STW*	*AGROVOC*	*TGN*	*UNESCO*	*ODT*	*SSWT*	*GBA-GU*	*GBA-GTS*	*GBA-L*	*GBA-LU*	*CECCT*
*DATA-MODEL-CONSISTENCY-01 (!)**												
*DATA-MODEL-CONSISTENCY-02 (!)**												
*DATA-MODEL-CONSISTENCY-03 (!)**												

Table 42: Thesauri Evaluation - Data Model Consistency (1)

	Data Sets
Data Model Consistency	*EARTh* *GEMET* *Euro Voc* *SLD*
*DATA-MODEL-CONSISTENCY-01 (!)**	
*DATA-MODEL-CONSISTENCY-02 (!)**	
*DATA-MODEL-CONSISTENCY-03 (!)**	

Table 43: Thesauri Evaluation - Data Model Consistency (2)

	Data Sets											
Labeling and Documentation	*TheSoz*	*STW*	*AGROVOC*	*TGN*	*UNESCO*	*ODT*	*SSWT*	*GBA-GU*	*GBA-GTS*	*GBA-L*	*GBA-LU*	*CECCT*
*LABELING-AND-DOCUMENTATION-01**	8,426	11,508	19,829	1,110	✗	36	1,475	5	2	✓	107	486
*LABELING-AND-DOCUMENTATION-02**	>1	✗	>100	287	✗	✓	✓	✓	✓	✓	✓	✓
*LABELING-AND-DOCUMENTATION-03**	✓	✓	1	14,114	✗	✓	✓	1	✓	✓	1	✓
*LABELING-AND-DOCUMENTATION-04 (!)**												
*LABELING-AND-DOCUMENTATION-05**	✓	✓	4	✓	1	2	2	1	✓	✓	✓	7
*LABELING-AND-DOCUMENTATION-06**	975,340	✓	✓	2	✓	✓	✓	✓	✓	✓	✓	✓

Table 44: Thesauri Evaluation - Labeling and Documentation (1)

	Data Sets			
Labeling and Documentation	*EARTh*	*GEMET*	*EuroVoc*	*SLD*
*LABELING-AND-DOCUMENTATION-01**	264,687	✗	54,911	31,195
*LABELING-AND-DOCUMENTATION-02**	✗	✗	✗	✓
*LABELING-AND-DOCUMENTATION-03**	2	✗	55,556	31,195
*LABELING-AND-DOCUMENTATION-04 (!)**				
*LABELING-AND-DOCUMENTATION-05**	39	✗	✗	978
*LABELING-AND-DOCUMENTATION-06**	302	46,718	✓	✓

Table 45: Thesauri Evaluation - Labeling and Documentation (2)

	Data Sets											
Structure	*TheSoz*	*STW*	*AGROVOC*	*TGN*	*UNESCO*	*ODT*	*SSWT*	*GBA-GU*	*GBA-GTS*	*GBA-L*	*GBA-LU*	*CECCT*
*STRUCTURE-01***	1	1,074	✓	✓	1	5	1	✓	✓	✓	✓	✓
*STRUCTURE-02 (!)**												
*STRUCTURE-03***	✓	✓	✓	✓	84	✓	✓	✓	✓	✓	✓	✓
*STRUCTURE-04**	2,906	8,046	726	✓	3,840	12	124	84	256	68	22	2,422
*STRUCTURE-05**	✓	✓	✓	✓	✗	90	5,150	✓	✓	✓	✓	9,864
*STRUCTURE-06**	1,457	37	✓	✓	✗	✓	4	1	1	64	✓	136
*STRUCTURE-07***	40	5,370	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓
*STRUCTURE-08 (!)****												
*STRUCTURE-09**	7,897	19,844	99	✓	552	2	16	26	✓	✓	✓	82
*STRUCTURE-10***	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Table 46: Thesauri Evaluation - Structure (1)

	Data Sets			
Structure	*EARTh*	*GEMET*	*Euro Voc*	*SLD*
*STRUCTURE-01***	18,240	✗	55,757	31,195
*STRUCTURE-02 (!)**				
*STRUCTURE-03***	39	4,244	✓	✓
*STRUCTURE-04**	11,286	74	✓	✓
*STRUCTURE-05**	✓	✗	✓	✓
*STRUCTURE-06**	239,346	✗	13,876	✓
*STRUCTURE-07***	110,015	✗	366,155	155,975
*STRUCTURE-08 (!)****				
*STRUCTURE-09**	107,195	32	✓	✓
*STRUCTURE-10***	27	2,122	✓	✓

Table 47: Thesauri Evaluation - Structure (2)

	Data Sets										
Language Tag Cardinality	*TheSoz*	*STW*	*AGROVOC*	*TGN UNESCO ODT*	*SSWT*	*GBA-GU*	*GBA-GTS*	*GBA-L*	*GBA-LU*	*CECCT*	
*LANGUAGE-TAG-CARDINALITY-01***	9,435	13,468	98,894	✓	541	10,147	5,117	2,061	1,742	2,272	15,550
*LANGUAGE-TAG-CARDINALITY-02**	8,222	36,936	✗	✓	265	3,627	2,212	635	631	1,253	9,607
*LANGUAGE-TAG-CARDINALITY-03**	8,222	✓	135	✓	✓	✓	✓	✓	✓	✓	✓
*LANGUAGE-TAG-CARDINALITY-04**	✓	476	✗	50	✓	✓	✓	✓	✓	✓	✓

Table 48: Thesauri Evaluation - Language Tag Cardinality (1)

	Data Sets			
Language Tag Cardinality	*EARTh*	*GEMET*	*EuroVoc*	*SLD*
*LANGUAGE-TAG-CARDINALITY-01***	✗	2,318,895	✗	30,781
*LANGUAGE-TAG-CARDINALITY-02**	✗	✗	✗	✗
*LANGUAGE-TAG-CARDINALITY-03**	224,206	✗	✗	31,195
*LANGUAGE-TAG-CARDINALITY-04**	✗	✗	✓	✓

Table 49: Thesauri Evaluation - Language Tag Cardinality (2)

	Data Sets									
Constraints	*TheSoz*	*STW*	*AGROVOC*	*TGN*	*UNESCO*	*ODT*	*SSWT*	*GBA-GU*	*GBA-GTS*	*GBA-L*	*GBA-LU*	*CECCT*
*PROPERTY-DOMAIN-01 (!)****												
*PROPERTY-RANGES-01 (!)****												
*DISJOINT-PROPERTIES-01 (!)****												
*DISJOINT-PROPERTIES-02 (!)****												
*DISJOINT-CLASSES-01 (!)****												
*EQUIVALENT-PROPERTIES-01 (!)**												
*UNIVERSAL-QUANTIFICATIONS-01 (!)****												
*CONTEXT-SPECIFIC-VALID-CLASSES-01 (!)**												
*CONTEXT-SPECIFIC-VALID-PROPERTIES-01 (!)**												
*RECOMMENDED-PROPERTIES-01 (!)**												
*VOCABULARY-01 (!)****												
*HTTP-URI-SCHEME-VIOLATION (!)****												

Table 50: Thesauri Evaluation - Constraints (1)

Constraints	Data Sets			
	EARTh	*GEMET*	*Euro Voc*	*SLD*
PROPERTY-DOMAIN-01 (!)***				
PROPERTY-RANGES-01 (!)***				
DISJOINT-PROPERTIES-01 (!)***				
DISJOINT-PROPERTIES-02 (!)***				
DISJOINT-CLASSES-01 (!)***				
EQUIVALENT-PROPERTIES-01 (!)*				
UNIVERSAL-QUANTIFICATIONS-01 (!)***				
CONTEXT-SPECIFIC-VALID-CLASSES-01 (!)*				
CONTEXT-SPECIFIC-VALID-PROPERTIES-01 (!)*				
RECOMMENDED-PROPERTIES-01 (!)*				
VOCABULARY-01 (!)***				
HTTP-URI-SCHEME-VIOLATION (!)***				

Table 51: Thesauri Evaluation - Constraints (2)

5 Rectangular Data (PHDD)

Thomas: ToDo

5.1 Evaluation Results

5.2 Data Sets Overview

5.3 Detailed Evaluation

6 Statistical Classifications (XKOS)

Thomas: ToDo

6.1 Evaluation Results

6.2 Data Sets Overview

Abbr.	Statistical Classifications
NAF	*Nomenclature d'activités française*[38]
PCS	*Nomenclature des Professions et Catégories Socioprofessionnelles*[39]
CJ	*Nomenclature des catégories juridiques*[40]
ISIC	[41]
ISCO	[42]

Table 52: Statistical Classifications Abbreviations

- *Nomenclature d'activités française (NAF)*: French classification expressed in XKOS. the French refinement of the NACE, because it has explanatory notes.
- *Nomenclature des Professions et Catégories Socioprofessionnelles (PCS)*: French classification expressed in XKOS.
- *Nomenclature des catégories juridiques (CJ)*: French classification expressed in XKOS.
- *ISIC*: has explanatory notes too.
- *ISCO*

[38] http://rdf.insee.fr/codes/index.html
[39] http://rdf.insee.fr/codes/index.html
[40] http://rdf.insee.fr/codes/index.html
[41]
[42]

	Counts
Data Sets	**triples**
NAF	
PCS	
CJ	
ISIC	
ISCO	
Total	

Table 53: Statistical Classifications Overview

6.3 Detailed Evaluation

References