

Abstract:
 Associations, which are one of the key ingredients of human intelligence and thinking, are not easily accessible to the Semantic Web community. High quality RDF datasets of this kind are missing.
 In this work we generate such a dataset by transforming 788 K free-text associations of the Edinburgh Associative Thesaurus (EAT) into RDF. Furthermore, we provide a verified mapping of strong textual associations from EAT to DBpedia Entities with the help of a semi-automatic mapping approach.
 Both generated datasets are made publicly available and can be used as a benchmark for cross-type link prediction and pattern learning.

Edinburgh Associative Thesaurus (EAT) as RDF and DBpedia Mapping

15

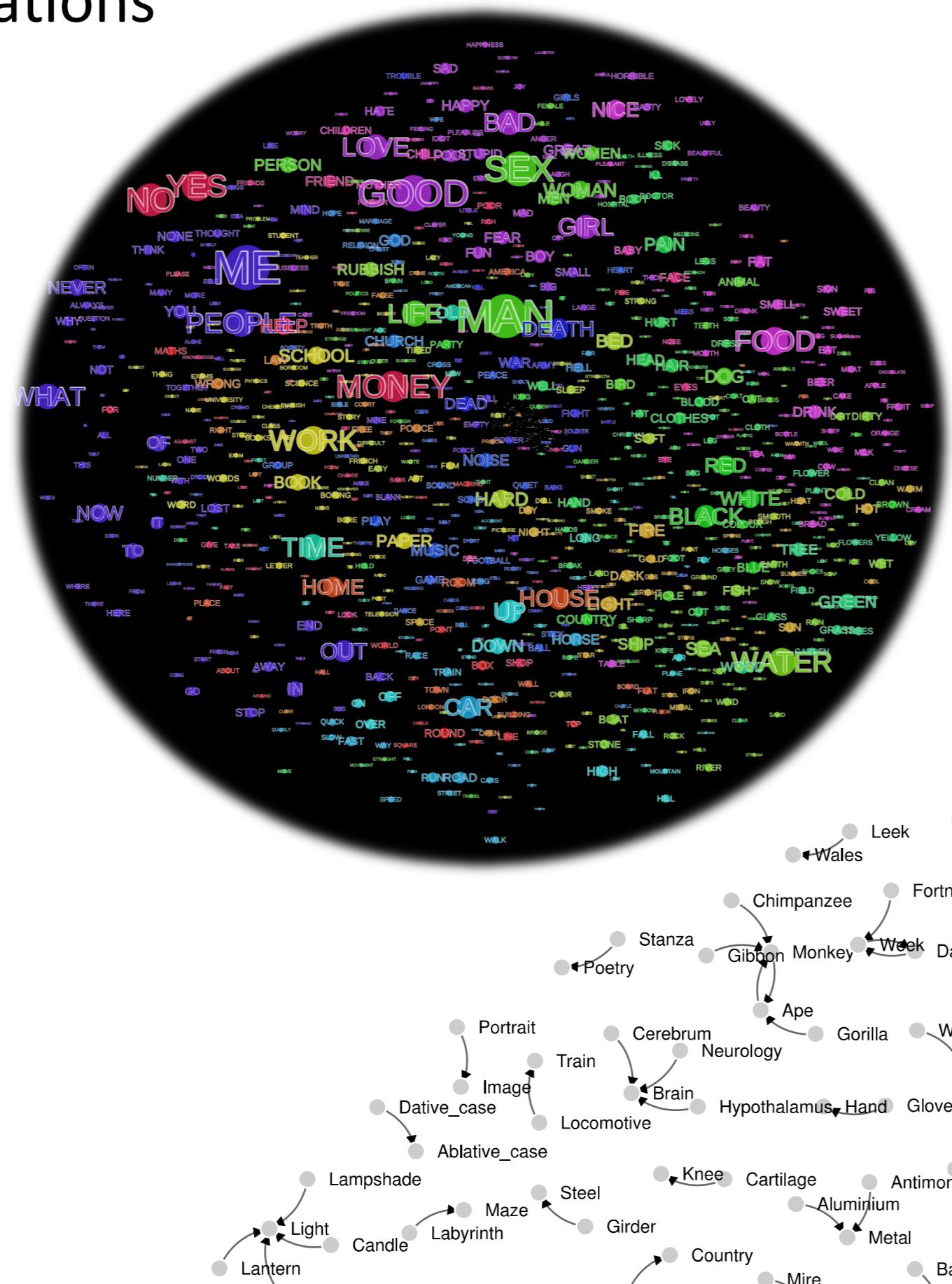
Jörn Hees, Rouven Bauer, Joachim Folz, Damian Borth, Andreas Dengel

EAT: G. Kiss, C. Armstrong, R. Milroy, J. Piper (1973)

- Association corpus
 - For each stimulus asked 100 people for a response
 - Strong responses became stimuli of next round
- 790 K raw associations (free text)
- Graph: ($|V| = 23\text{K}$, $|E| = 325\text{K}$)
- 5000 strong associations (>19x)
 - 167.4 K raw associations

Example Associations:

dog	cat
cat	57
collar	5
bark	3
leg	2
tom	2
man	woman
woman	66
strong	3
human	2
hole	2
boy	2
bank	money
money	25
book	7
account	7
manager	6
clerk	5
gold	6
wealth	5



Resulting Top Response Nodes

Response	Count	Response	Count
db:Person	10	db:User	9
db:Bird	15	db:Army	8
db:Horse	14	db:Beer	8
db:Automobile	13	db:Death	7
db:Flower	12	db:Fish	7
db:Music	12	db:Bed	7
db:Treasure	11	db:Ship	7
db:Sea	11	db:Red	6
db:Dog	9	db:Green	6
db:Food	9	db:Hair	6

Mapping Verification:

- Web App
 - Allows quick manual verification
- Test group (13 people)
 - Valid: 3 people say "correct"
 - Invalid: Everything else (for now)

➤ 790 valid semantic associations

- 25.5 K raw associations

Please verify the mappings of these associated words to Wikipedia

ARITHMETIC (stimulus)

MATHS (response)

Are both pages correct?

(same meaning as the stimulus)

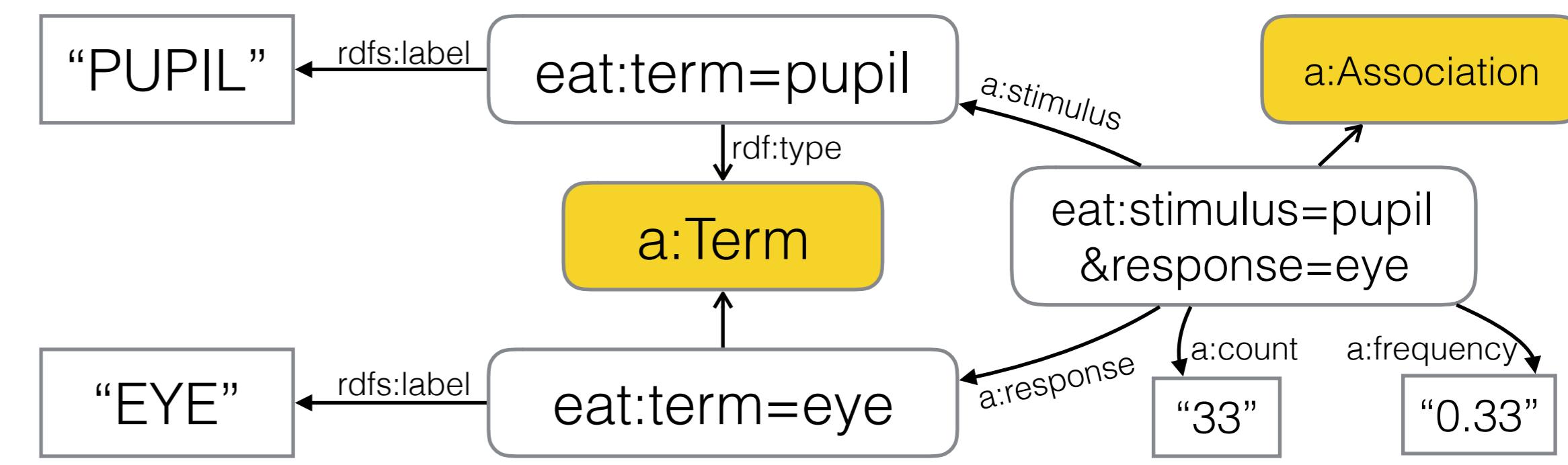
Arithmetic
Arithmetic or arithmetics from the Greek word ἀριθμός, arithmeticus "number" is the oldest and most elementary branch of mathematics. It consists in the study of numbers, especially the properties of the traditional operations between them – addition, subtraction, multiplication, division, powers, extraction of roots, etc. Mathematics is also the study of the year of number theory and number theory is considered to be one of the top-level divisions of modern mathematics, along with algebra, geometry, and analysis. The terms arithmetic and higher arithmetic were used until the beginning of 20th century as synonyms for number theory, and sometimes, still used to refer to a wider part of number theory.

(same meaning as the response)

Mathematics
Mathematics is the study of topics such as quantity (numbers), structure, space, and change. There is a range of views among mathematicians and philosophers as to the exact scope and definition of mathematics. Mathematicians seek out patterns and use them to formulate new conjectures. Mathematicians explore the consequences of these conjectures by mathematical deduction. When mathematical structures are well models of real phenomena, then mathematical reasoning can provide insight or predictions about nature. Through the use of abstraction and logic, mathematics developed from counting, calculation, measurement, and

EAT as RDF:

- 1.6 M triples



Automatic Mapping Approach to DBpedia:

- Lookup + Scoring on Wikipedia

• Scoring:

- Composite phrases ("port" - "wine")
- Word types non-nouns ("unbound" - "free")
- Reflexive mappings ("child" - "children")
- Plural words ("thumbs" - "fingers")
- Disambiguation pages ("pod" - "pea")

• Results:

- 1066 mapping candidates
- Manual verification

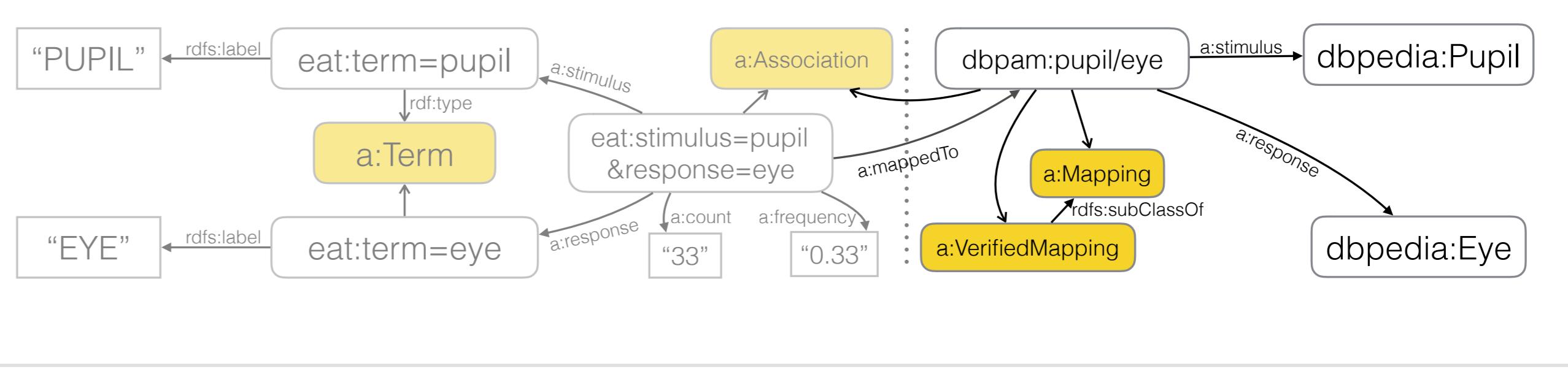
Node Degrees in Dbpedia

Vanilla:	Extended (Wikilinks+):
Node	Degree
db:Animal	237855
db:Insect	118889
db:Species	9802
db:France	231700
db:India	85286
db:Plant	79062
db:Italy	196686
db:Village	5076
db:Town	43730
db:Soiland	27607
db:Beetle	83109
db:Scotland	73312
db:Bird	25933
db:Switzerland	19874
db:Paris	66504
db:Germany	149344
db:Paris	17362
db:Wales	53008
db:Bird	14603
db:Town	50332
db:Ireland	13301
db:Marriage	40592
db:Religion	38673
db:Wales	10299
db:Fly	38532
db:Mayor	9812
db:School	32824
db:Reptile	9595
db:Novel	32193

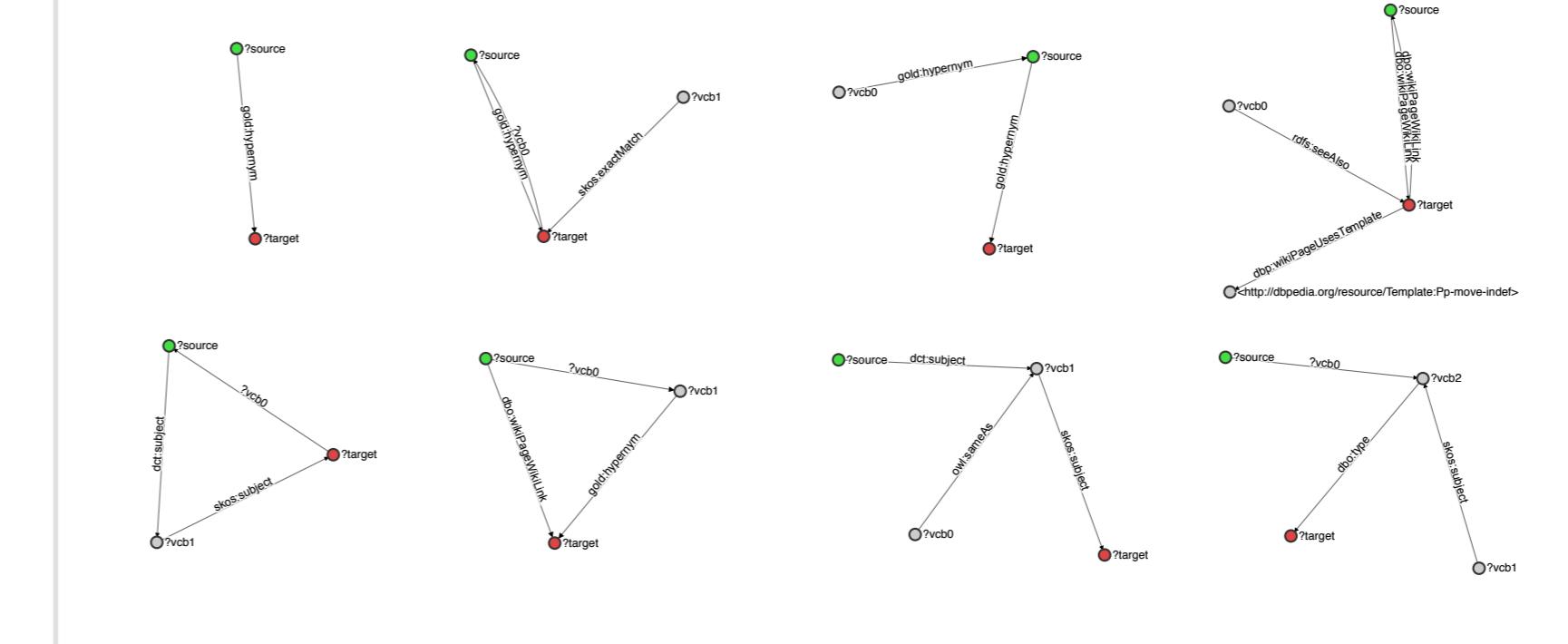
Semantic Associations:

• Results:

- 727 unique associations
- between DBpedia Entities



Future Work: Graph Pattern Learning



w3id.org/associations



joern.hees@dfki.de

DFKI German Research Center for Artificial Intelligence KAISERSLAUTERN