

Tiessen, Jan (2007): Die Resultate im Blick?
net/Döhler, Marian (Hrsg.): Agencies in W
Tonderf, Karin/Bahnmüller, Reinhard/Klages,
instrument. Anwendungspraxis, Probleme
sigma.
Touraine, Alain (1984): Le retour de l'acteur; e
Treiber, Hubert (1984): Warum man nicht die
Mikroskop den ganzen Elefanten zu schen.



Analyzing the Evolution of Linked Vocabularies

Mohammad Abdel-Qader^{1,2}, Iacopo Vagliano¹, and Ansgar Scherp³

¹ ZBW – Leibniz Information Centre for Economics

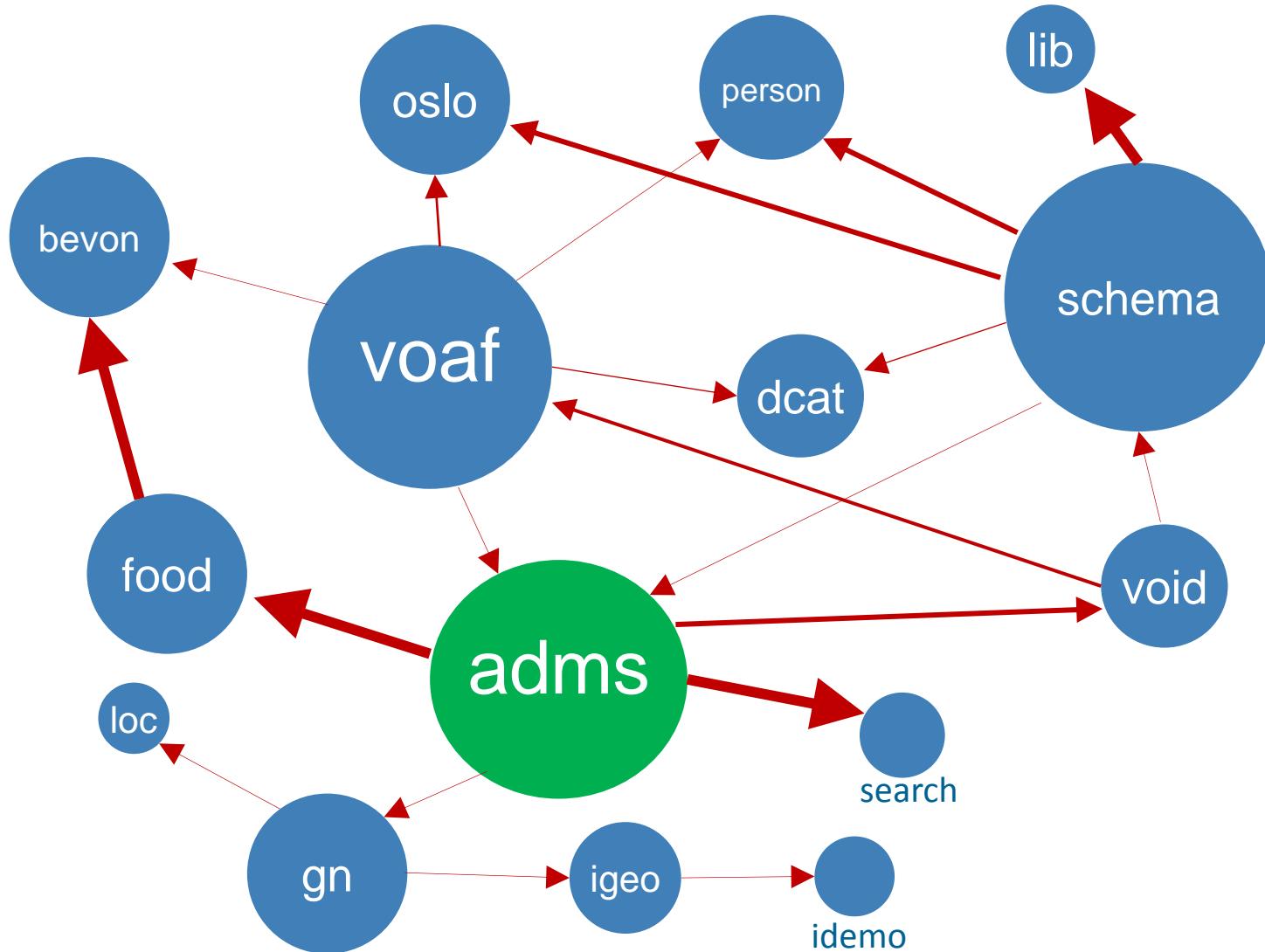
² Christian-Albrechts-University in Kiel, Germany

³ University of Essex, United Kingdom

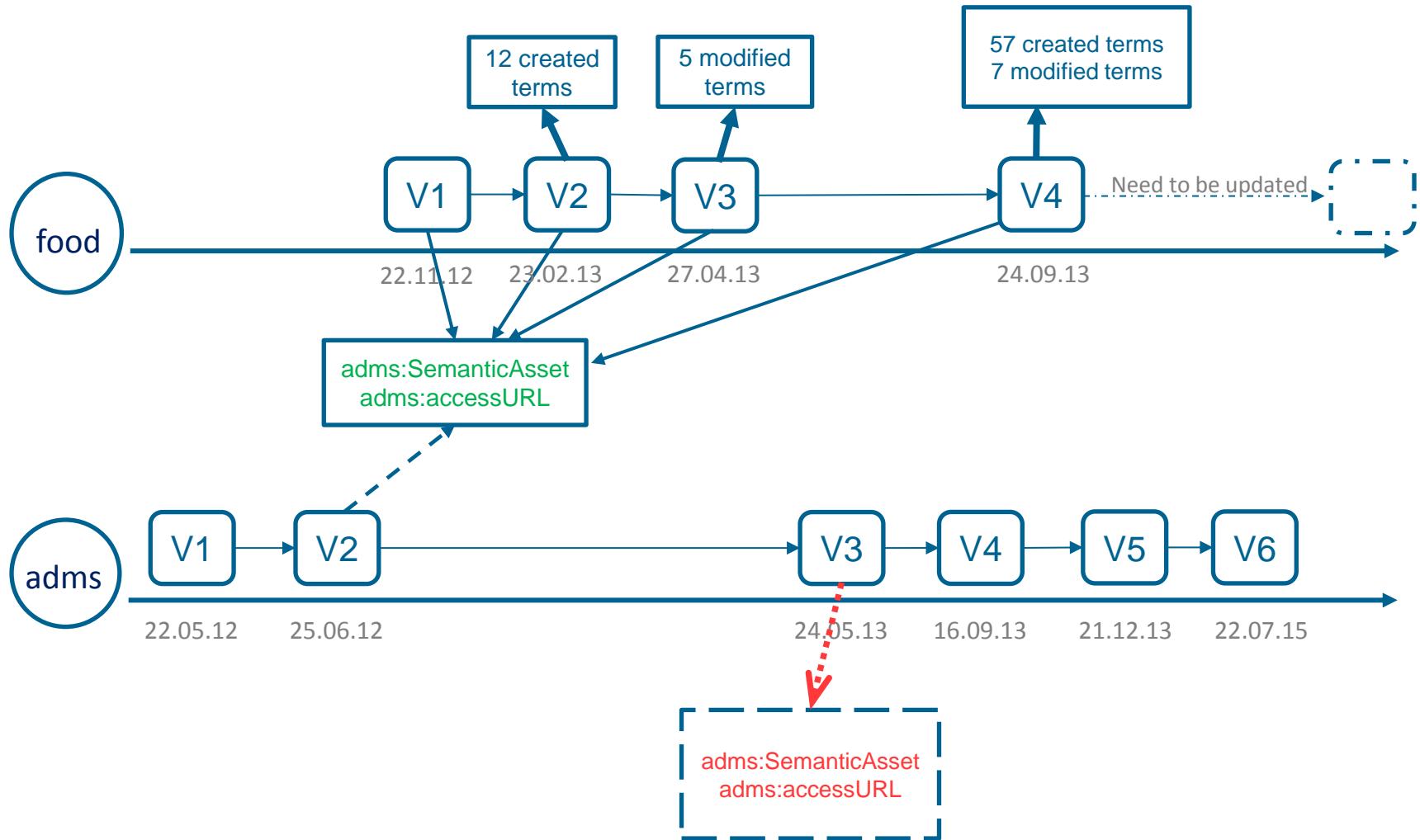
June 14th , 2019, 19th ICWE 2019
11.06.2019- 14.06.2019, Daejeon, Korea

www.moving-project.eu

Motivation example 1



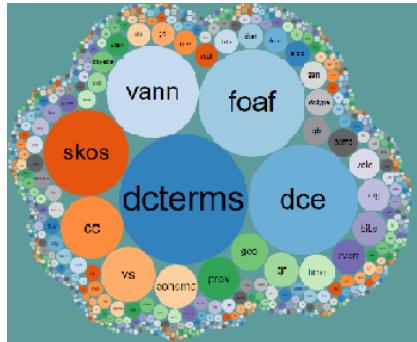
Motivation example 2



Research Questions

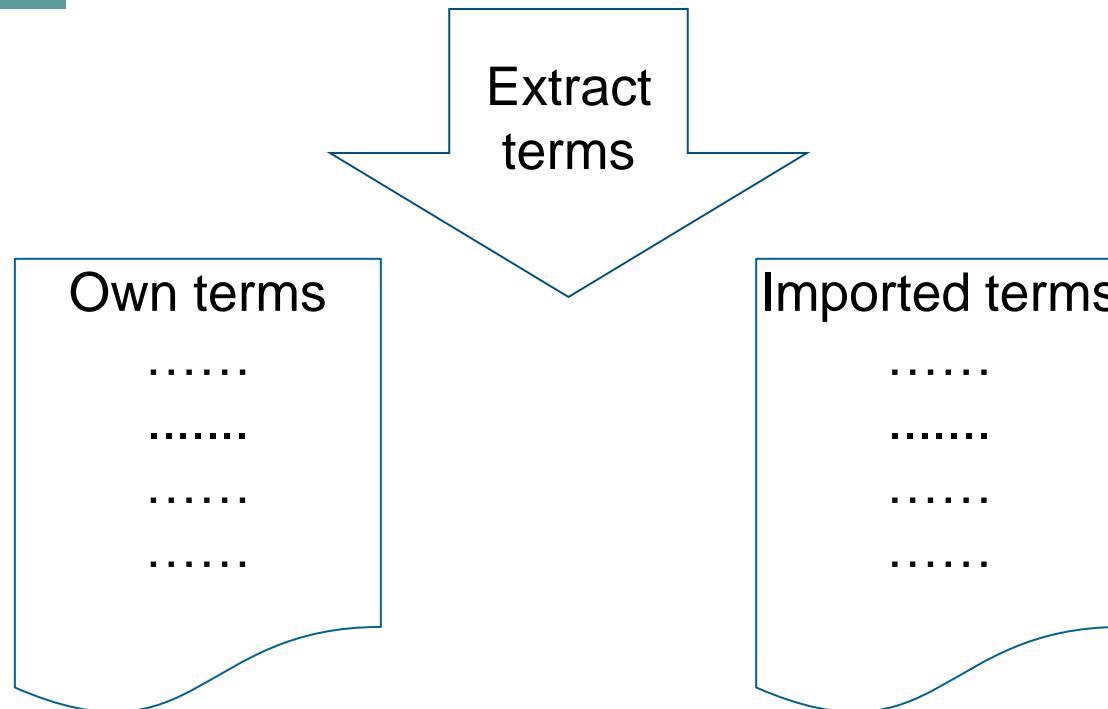
- RQ1: What is the state of the Network of Linked Vocabularies?
 - Size?
 - Density, average degree?
 - Central nodes?
- RQ2: How are vocabulary terms reused by other vocabularies?
 - How many?
 - Most recent ones?
 - Change impact?
- RQ3: How do ranking metrics, such as PageRank, HITS, and Centrality, change during the evolution of the Network of Linked Vocabularies?
 - Important nodes?
 - Reuse over time?

Analysis Methodology- Step1



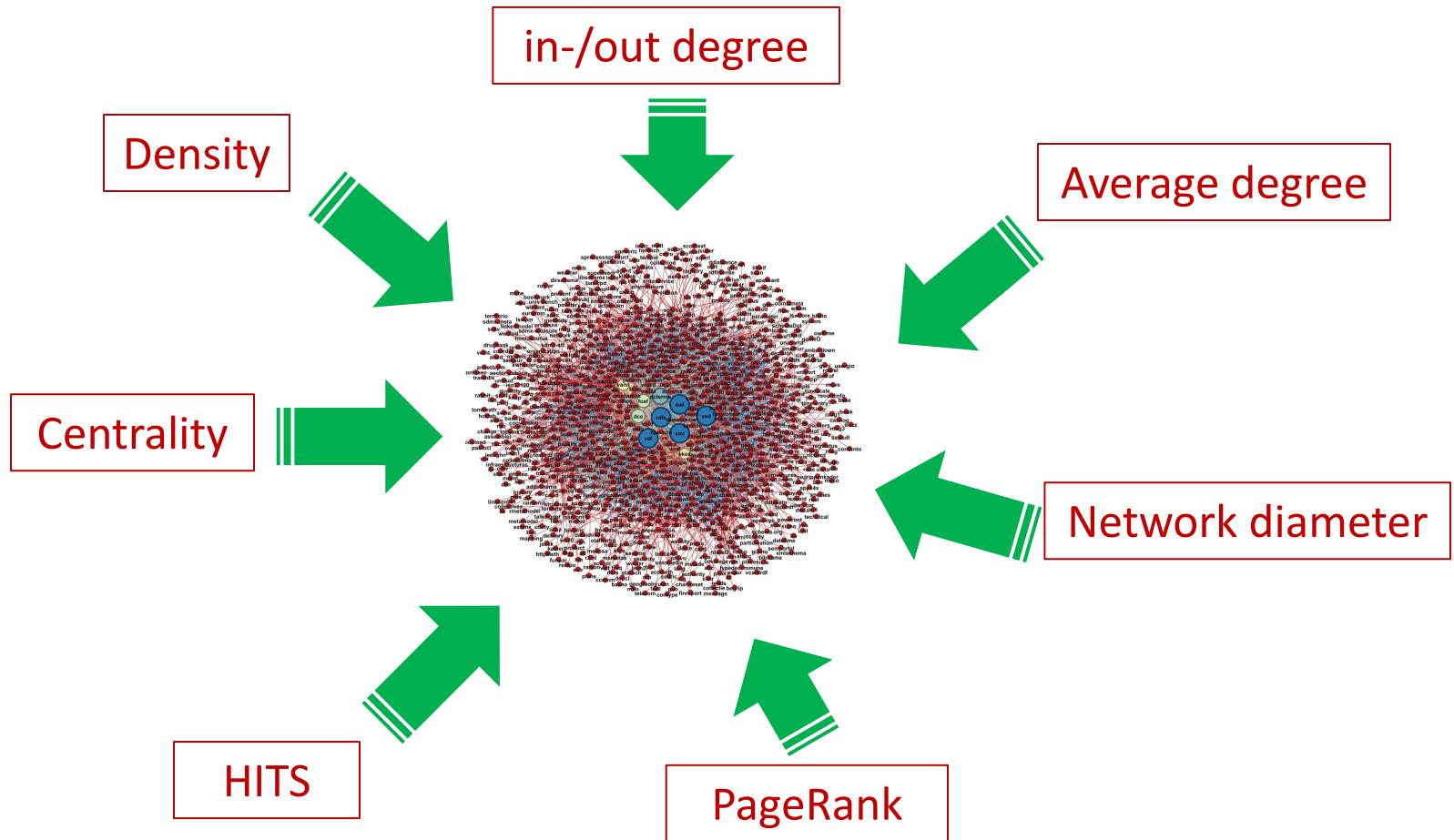
LOV- Linked Open Vocabulary¹

636 vocabularies until June 2018

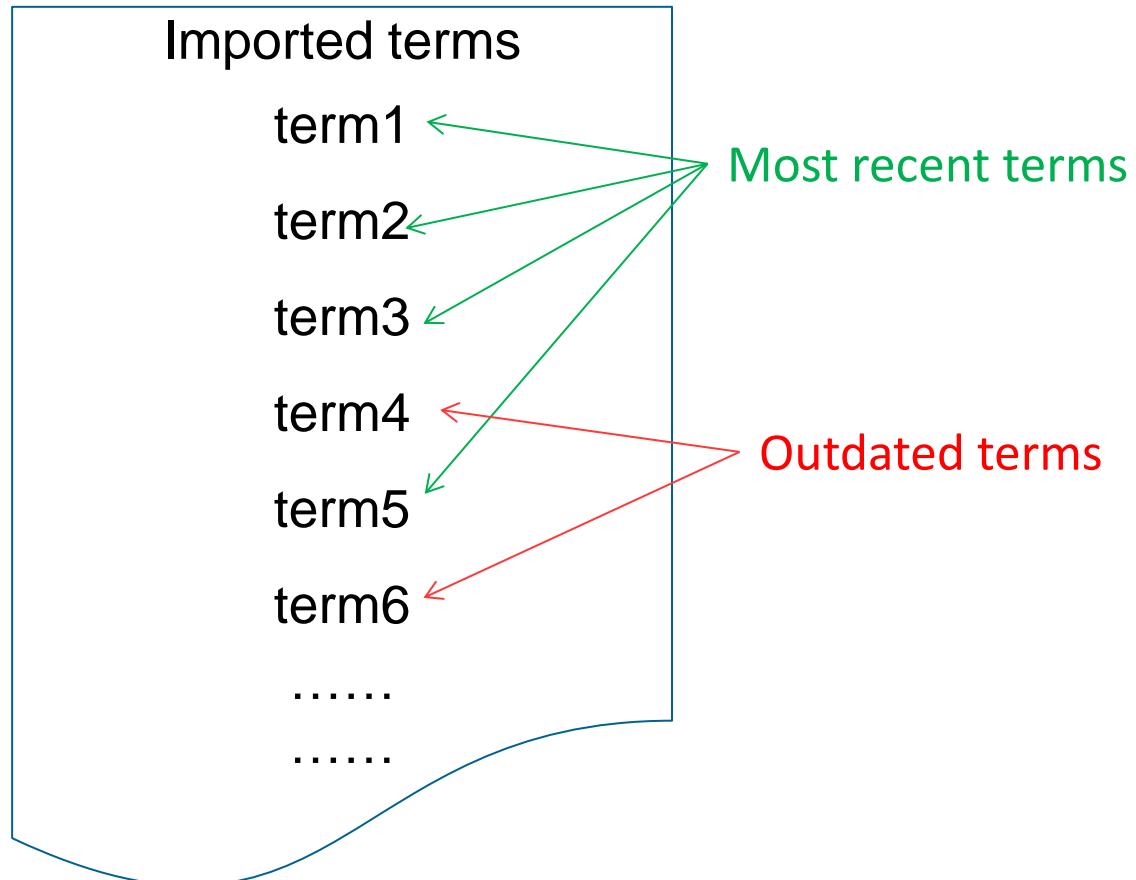


¹<http://lov.okfn.org/dataset/lov/>

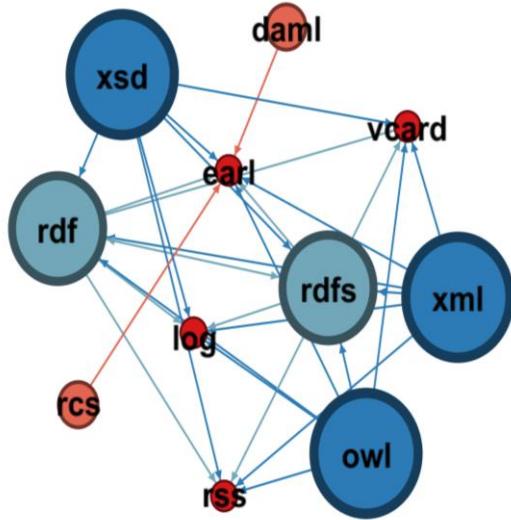
Analysis Methodology- Step2



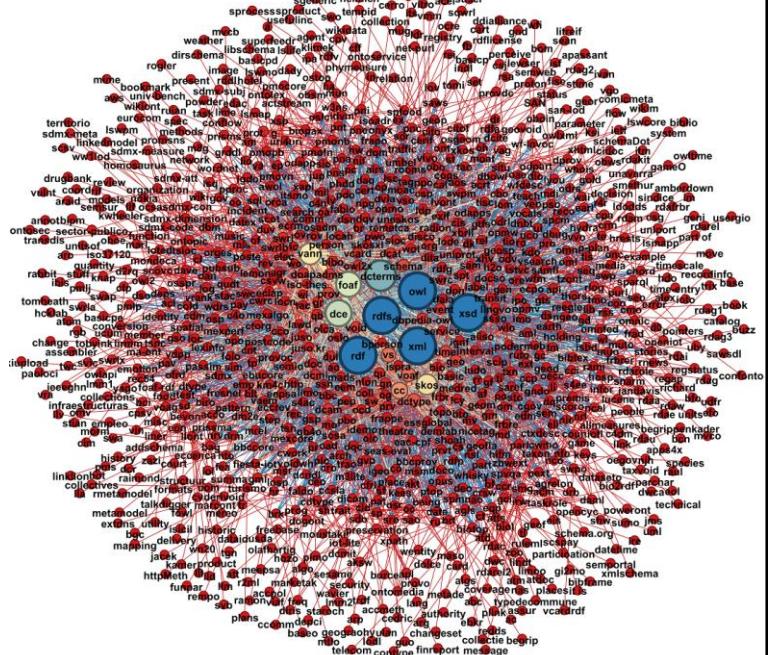
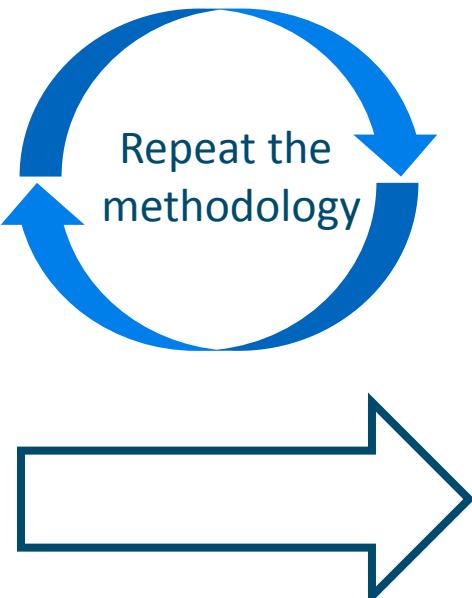
Analysis Methodology- Step3



Analysis Methodology- Step4



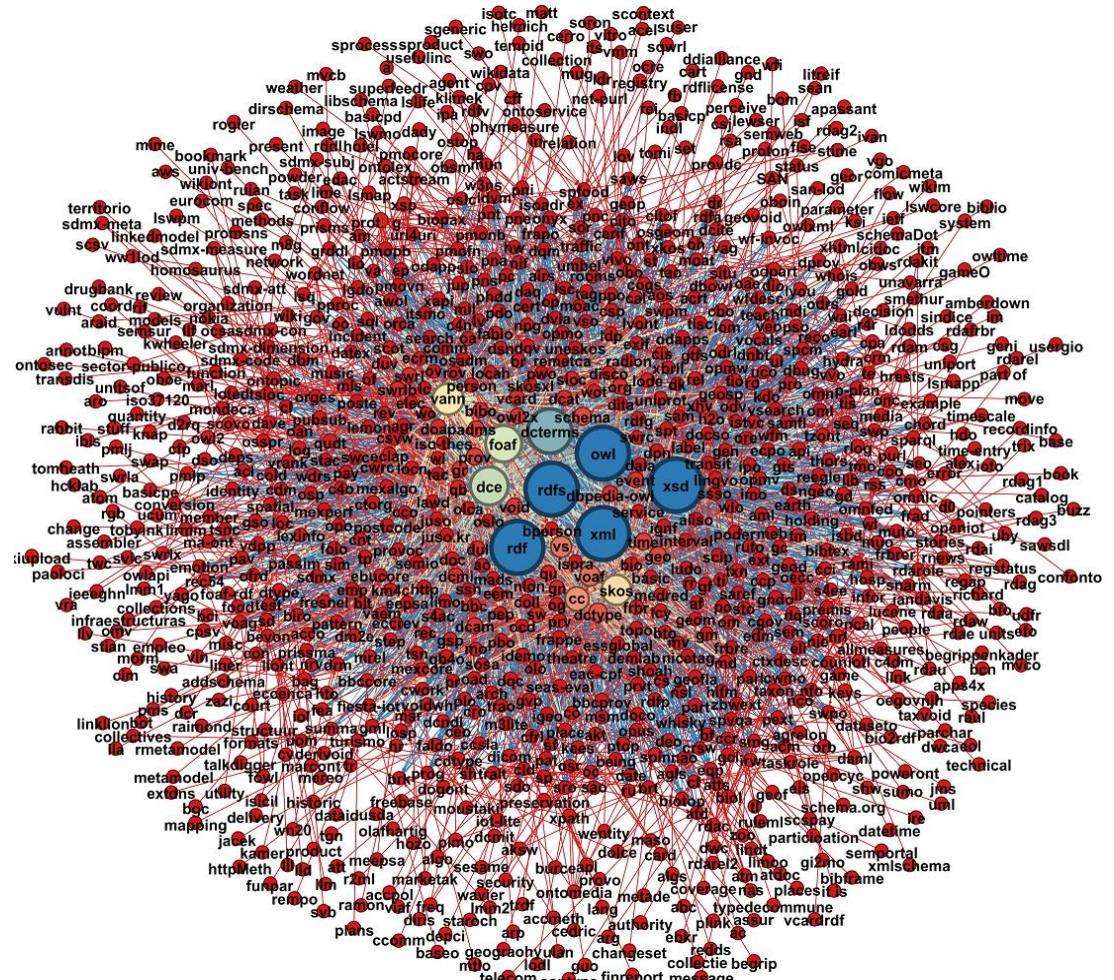
2001



2018

Results-RQ1 State of NeLO 2018

Measure	Value
Nodes	994
Edges	7046
Diameter	12
Density	0.007
Degree	7.089



The full results and scalable version of figures are available at: <https://sites.google.com/view/nelo-evolution>

Results-RQ1 State of NeLO 2018

Top-10 vocabularies for HITS (Hub and Authority) scores in 2018

Vocabulary	Authority	Hub
dcterms	0.305421	0.037978
dce	0.242374	0.037727
foaf	0.234664	0.044112
vann	0.184754	0.045030
skos	0.171827	0.034529
cc	0.113386	0.034723
vs	0.081972	0.040256
voaf	0.080739	0.045920
dtype	0.058152	0.037727
schema.org	0.046659	0.040364

Top-10 vocabularies for PageRank scores in 2018

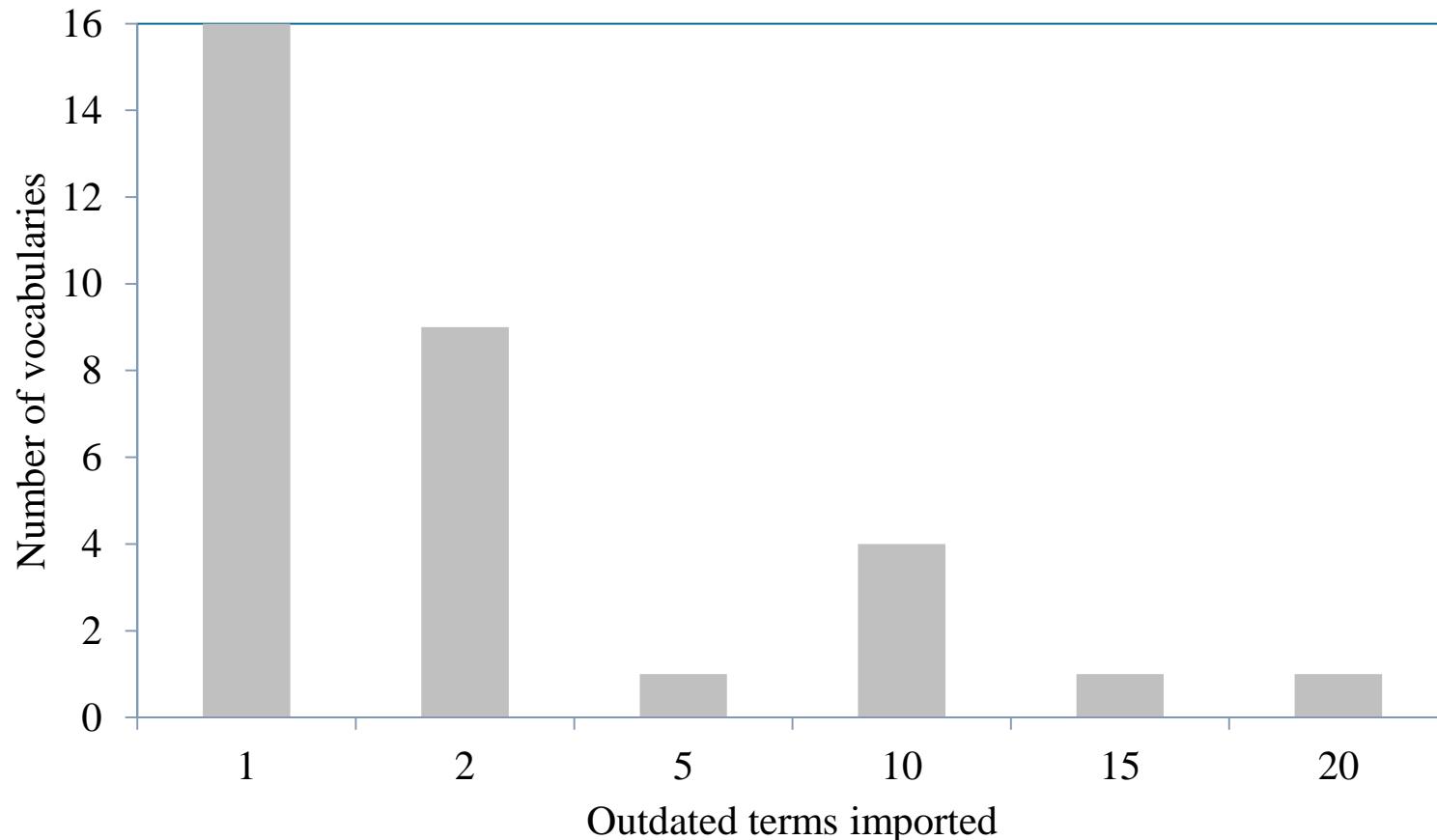
Vocabulary	PageRank
dce	0.045954
dcterms	0.027649
skos	0.017678
foaf	0.013986
dcam	0.009152
vann	0.009117
grddl	0.008740
dtype	0.005744
cc	0.005446
vs	0.005005

Results-RQ2 Reuse of vocabulary terms

Top-10 terms that are reused by other vocabularies in 2018.

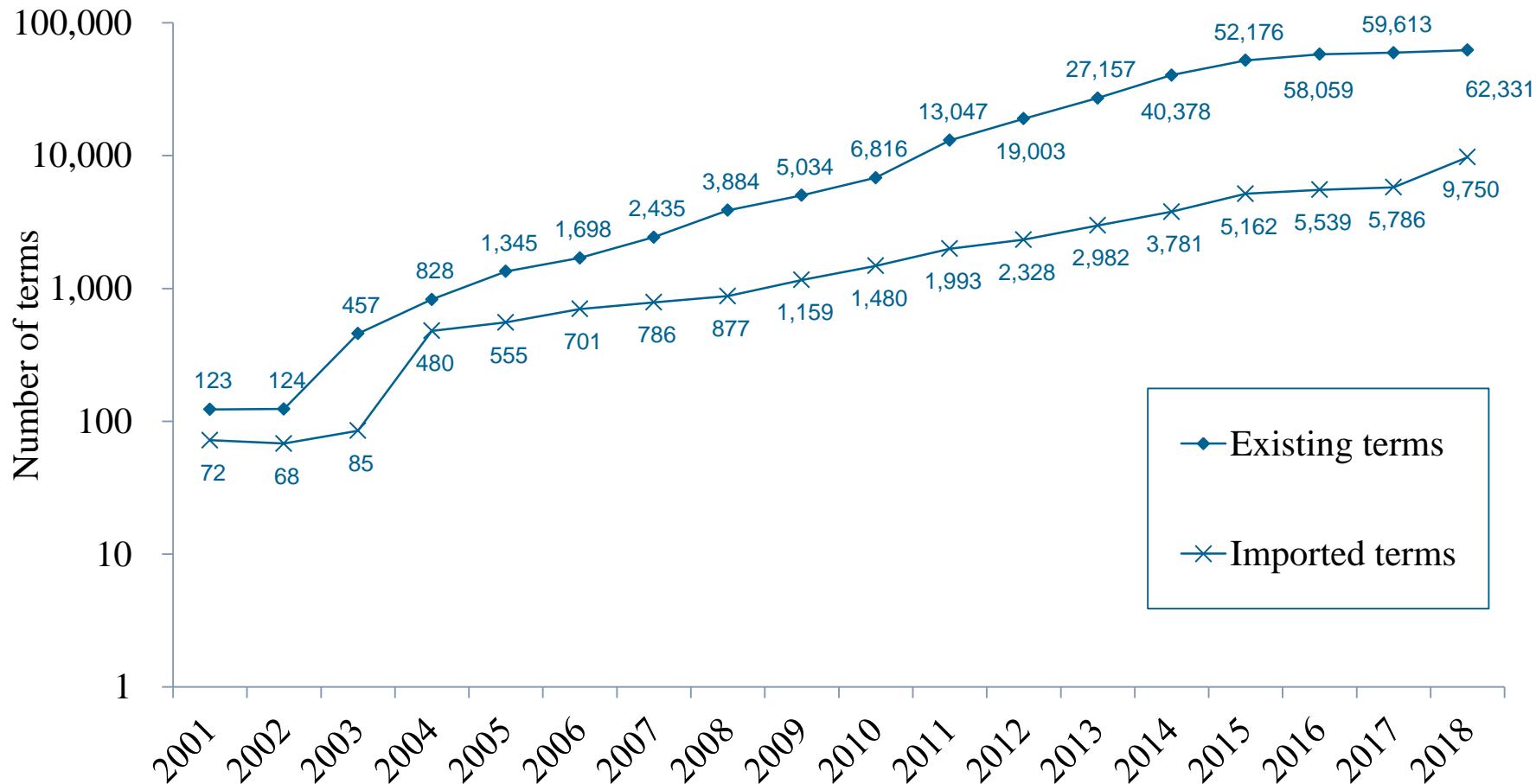
Term	Importing vocabularies.
dcterms:modified	281
dcterms:title	276
dce:title	266
dce:creator	263
vann:preferredNamespacePrefix	257
dcterms:description	249
vann:preferredNamespaceUri	241
foaf:Person	175
foaf:name	164
cc:license	122

Results-RQ2 Reuse of vocabulary terms



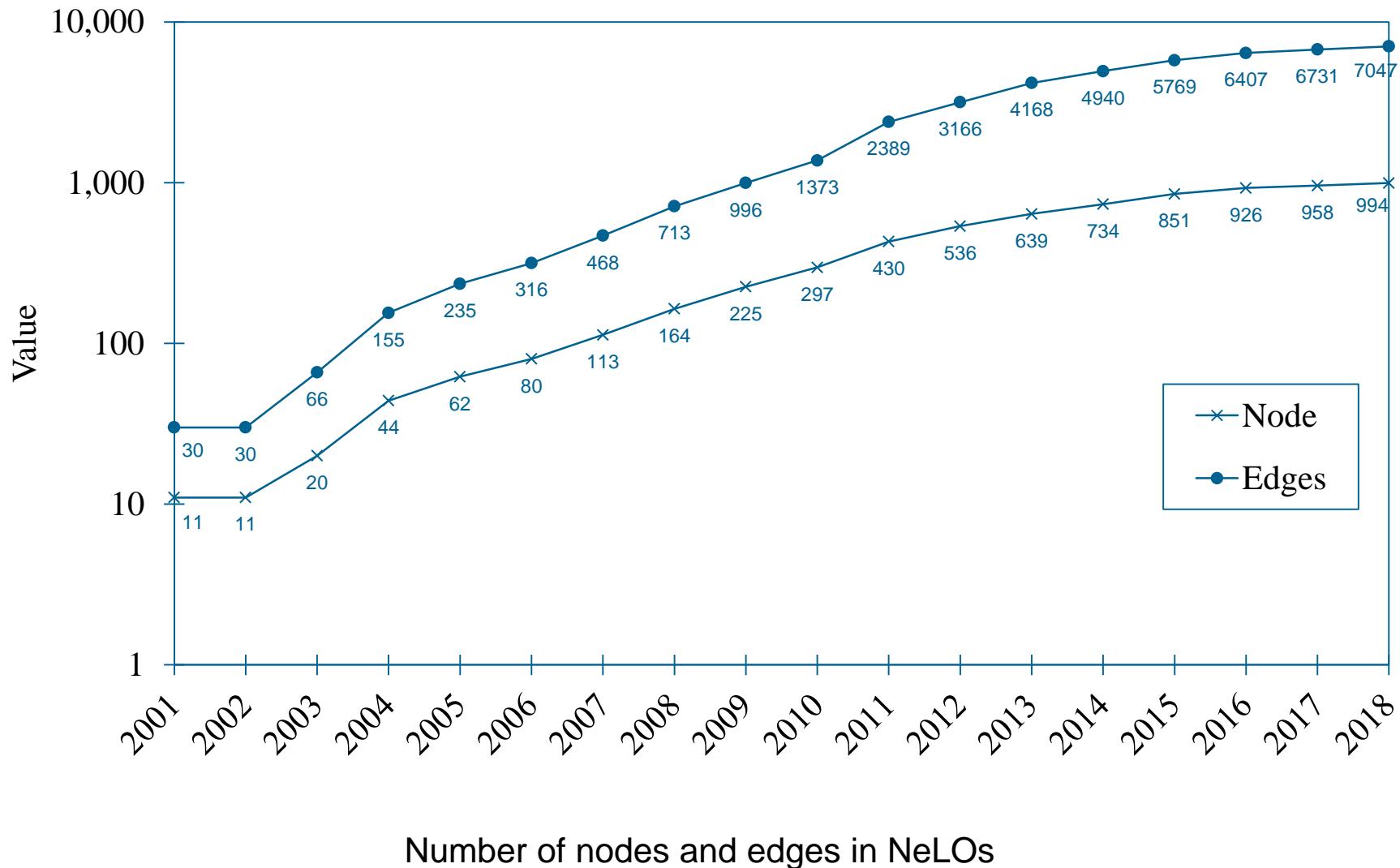
The number of vocabularies that reuse outdated terms by the number of outdated terms reused

Results-RQ2 Reuse of vocabulary terms

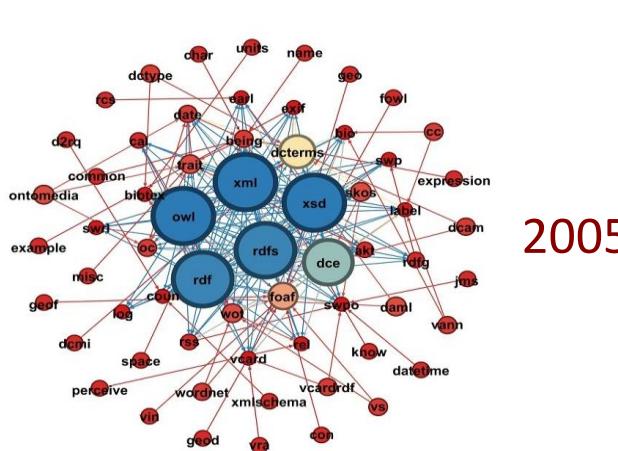


The number of existing terms and the reused terms from other vocabularies

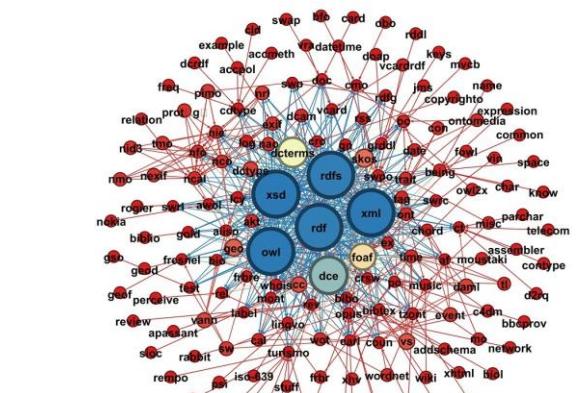
Results-RQ3 NeLO evolution



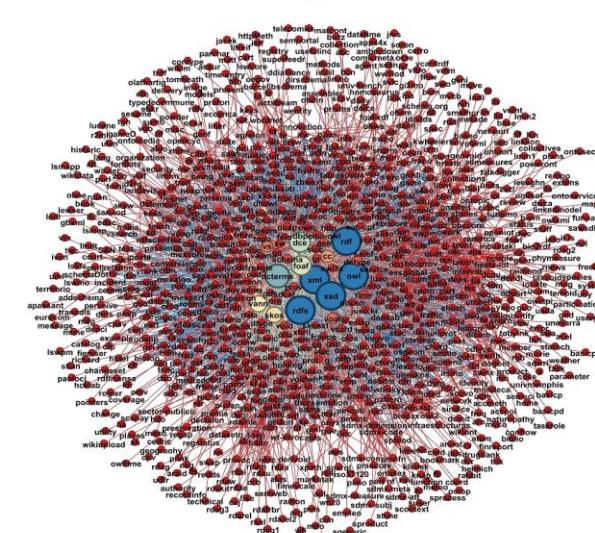
Results-RQ3 NeLO evolution



2005



2008

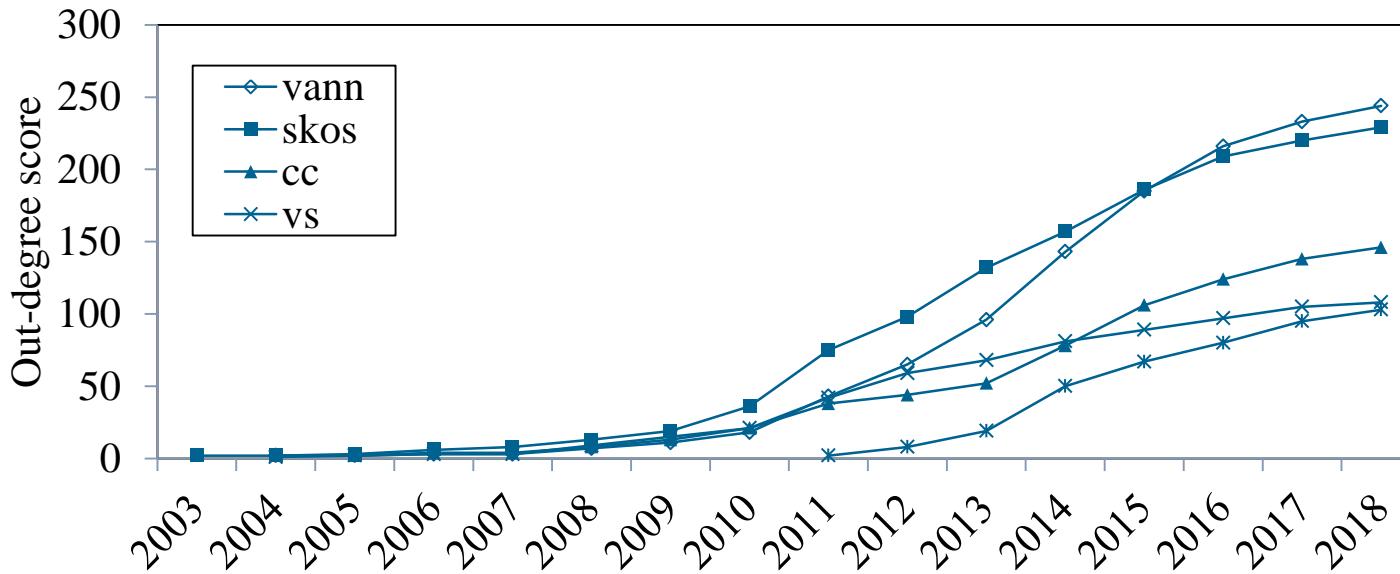
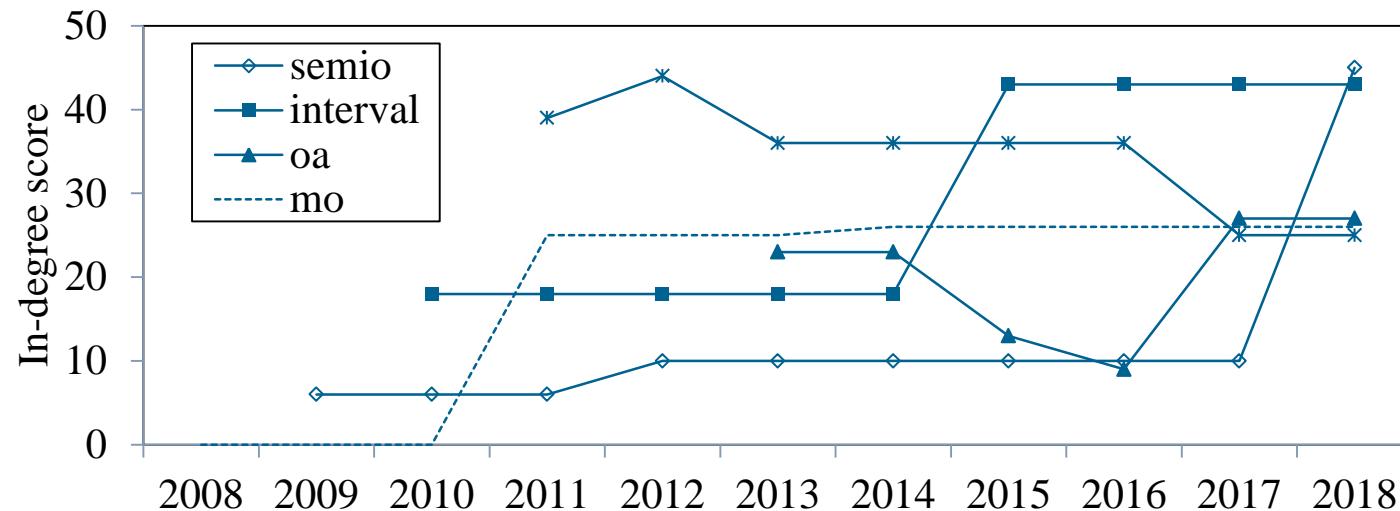


2017

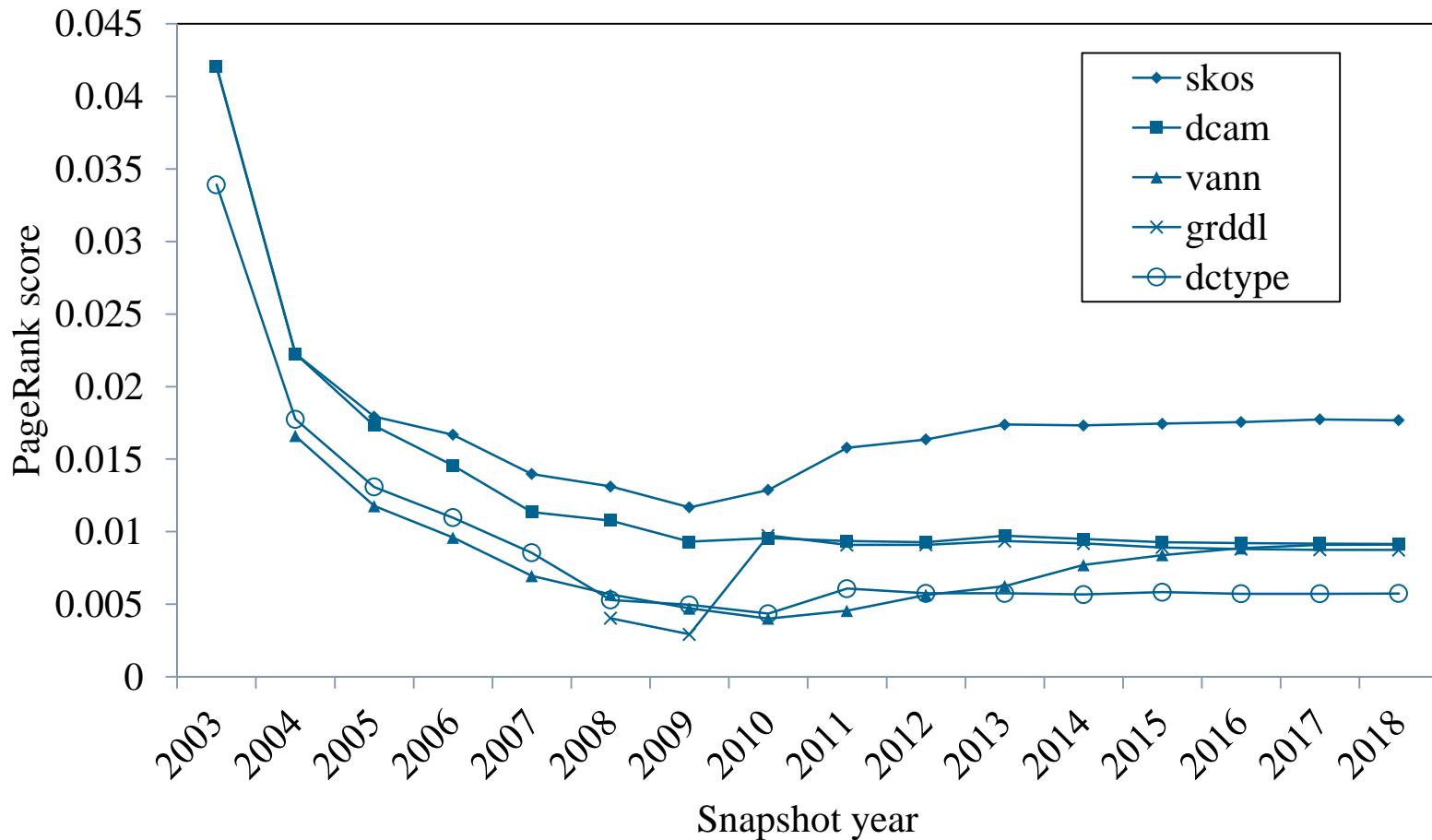
The full results and scalable version of figures are available at:

<https://sites.google.com/view/nelo-evolution>

Results-RQ3 NeLO evolution (In/out degree)

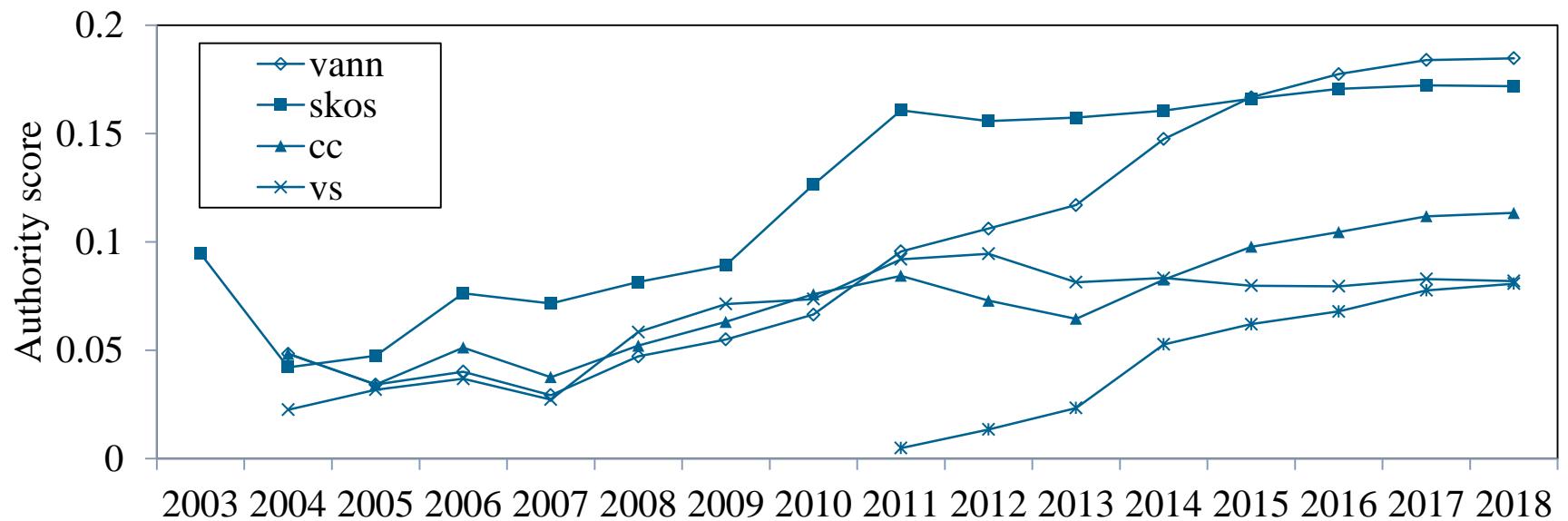
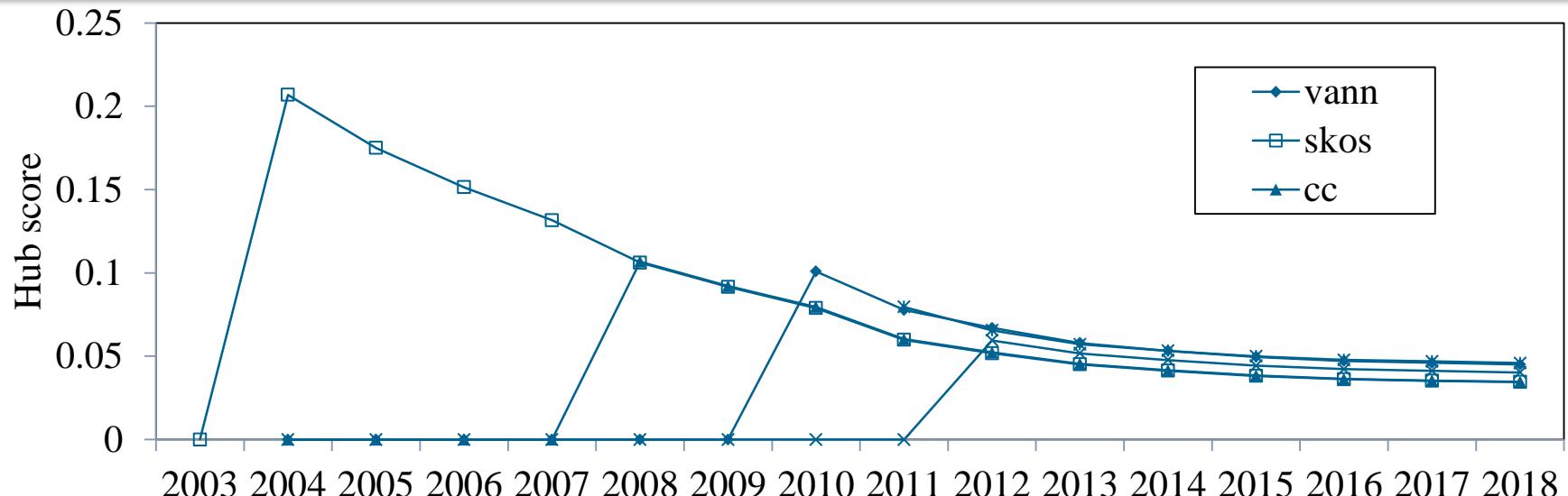


Results-RQ3 NeLO evolution



The PageRank scores for the top-five vocabularies for each year

Results-RQ3 NeLO evolution (HITS)



Conclusions

- From NeLO's 2018 density and average degree, we think that there is a need to increase the imports/exports relations between vocabularies.
- Only 16% of the existing terms are reused by the other vocabularies.
 - Using recommending systems can help to increase the reuse amount between vocabularies in order to avoid overlap and redundancy.
- The process of checking for the changes on other vocabularies is usually done manually.
- Many vocabularies are up-to-date in NeLO 2018.
 - 33 vocabularies are still using outdated terms.

Conclusions

- There is a lack of tools to notify ontology engineers about changes in the vocabularies.
- The dynamics of changes has slowed down after some fast evolution between 2001 and 2010.
- Based on the results, the reusing of terms was initially more common.
- Changes in the vocabularies with high PageRank and HITS scores affects many other vocabularies.

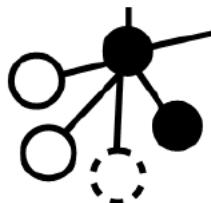
More results and scalable version of figures are available at:
<https://sites.google.com/view/nelo-evolution>

This work is supported by...



TraininG towards a society of data-saVvy inforMation prOfessionals to
enable open leadership INnovation

AND



Tiessen, Jan (2007): Die Resultate im Blick?
ner/Döhler, Marian (Hrsg.): Agencies in W
Tondorf, Karin/Bahnmüller, Reinhard/Klages,
instrument. Anwendungspraxis, Probleme
sigma.
Touraine, Alain (1984): Le retour de l'acteur: e
Treiber, Hubert (1984): Warum man nicht die
Mikroskop den ganzen Elefanten zu sehen.



Linked Open Citation Database
LOC-DB

References

- Abdel-Qader, M., Scherp, A., Vagliano, I.: Analyzing the evolution of vocabulary terms and their impact on the LOD Cloud. In: ESWC. Springer (2018)
- Cardoso, S.D., Pruski, C., Da Silveira, M., Lin, Y.C., Gro, A., Rahm, E., Reynaud-Delaitre, C.: Leveraging the impact of ontology evolution on semantic annotations. In: European Knowledge Acquisition Workshop. pp. 68-82. Springer (2016)
- Dividino, R., Gottron, T., Scherp, A.: Strategies for efficiently keeping local linked open data caches up-to-date. In: ISWC. pp. 356-373. Springer (2015)
- Dos Reis, J.C., Pruski, C., Da Silveira, M., Reynaud-Delaitre, C.: Understanding semantic mapping evolution by observing changes in biomedical ontologies. Journal of biomedical informatics 47, 71-82 (2014)
- Ghazvinian, A., Noy, N.F., Jonquet, C., Shah, N., Musen, M.A.: What four million mappings can tell you about two hundred ontologies. In: ISWC. pp. 229-242. Springer (2009)
- Gottron, T., Gottron, C.: Perplexity of index models over evolving linked data. In: ESWC. pp. 161-75. Springer (2014)
- Käfer, T., Abdelrahman, A., Umbrich, J., O'Byrne, P., Hogan, A.: Observing linked data dynamics. In: ESWC. pp. 213{227. Springer (2013)