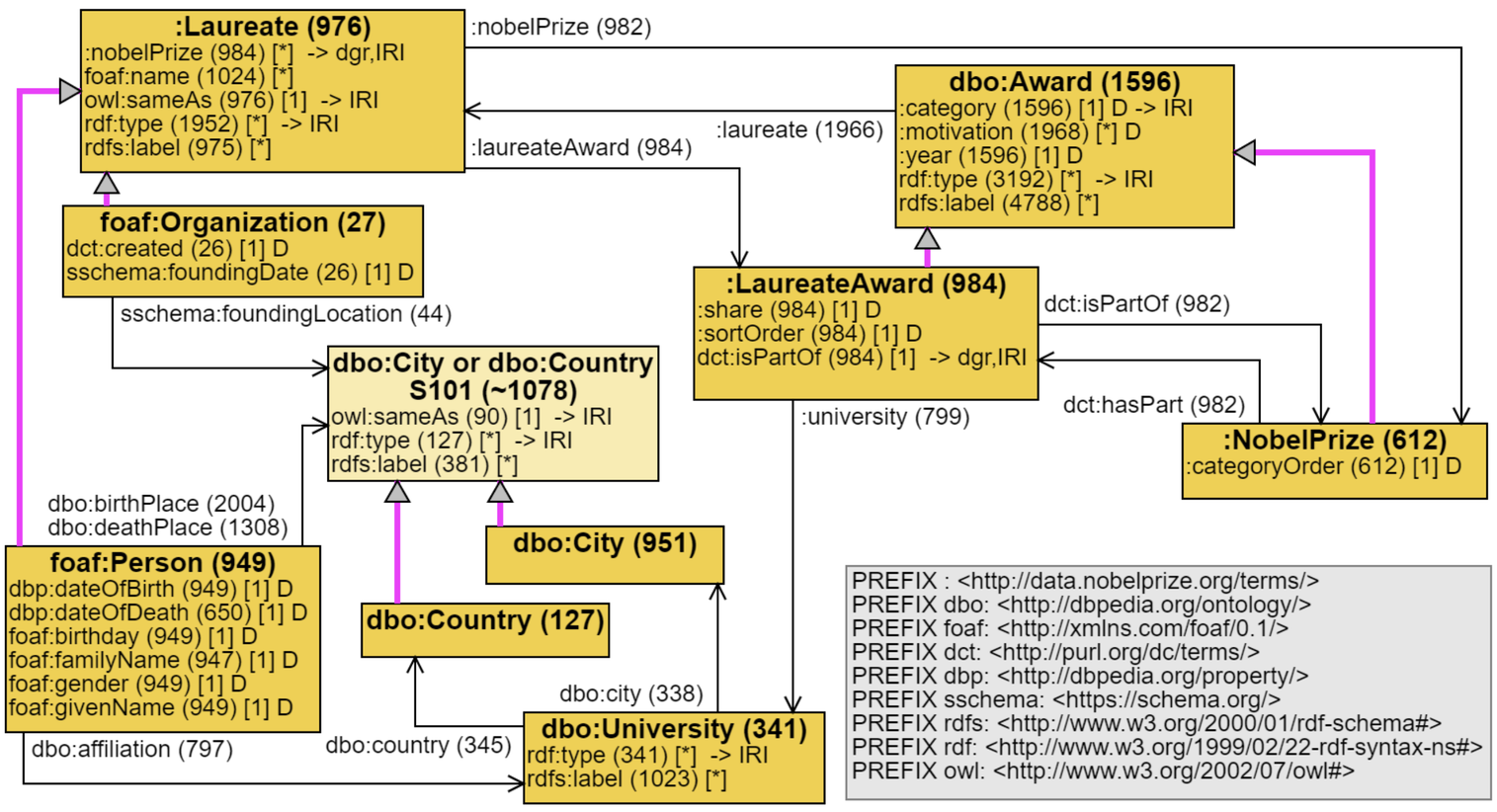
Schema Fragment Visualization to View Knowledge Graph Entities in Context

Sandra Siliņa, Lelde Lāce, Mikus Grasmanis, *Kārlis Čerāns*
Institute of Mathematics and Computer Science, University of Latvia
karlis.cerans@lumii.lv

ViziQuer: visual data schema exploration and visual queries

- ***Extract schema*** from a data endpoint, store in the tools' database
- ***Visualize the schema*** (in plain or ***summary*** form) or its fragment as a UML style diagram
- Small schemas give ***nice diagrams***
- Larger data schemas: schema diagrams unreadable or uncomputable
- How to select ***relevant fragments***?



Visual schema of a Nobel Prizes data set snapshot (original location: https://data.nobelprize.org/)

Schema size for some LOD data sets [1]

Endpoint URL	Triples	Cl.	Props.	Nodes	Lines
https://libris.kb.se/sparql	2695020210	540	939	225	467(0), 417(100)
https://sparql.nextprot.org/	2130123293	165	238	39	66(0), 65(100)
http://affymetrix.bio2rdf.org/sparql	1377023559	661	1328	132	307(0), 271(100)
https://ruian.linked.opendata.cz/sparql	870638775	85	200	43	53(0), 26(100)
http://data.bnf.fr/sparql	651506623	26	886	19	39(0)
http://kaiko.getalp.org/sparql	522998164	128	245	31	34(0), 28(100)
http://cr.eionet.europa.eu/sparql	482077457	272	2001	220	350(0), 308(100)
http://dati.isprambiente.it/sparql	385222839	122	383	86	135(0), 101(100)
http://dati.camera.it/sparql	322885735	92	283	61	119(0), 113(100)
http://data.allie.dbcls.jp/sparql	287461727	43	201	28	37(0)
http://datos.bne.es/sparql	258140051	16	329	15	32(0)
http://rdf.disgenet.org/sparql/	99381703	110	665	46	78(0), 48(100)
https://taxref.mnhn.fr/sparql	82745498	1913	608	58	96(0), 73(100)
http://opendata.aragon.es/sparql	70049160	206	1259	88	119(0), 103(100)
Muziekweb	37114240	30	59	16	20(0), 20(100)
http://premon.fbk.eu/sparql	32611819	95	146	46	42(0), 36(100)
http://geo.linkeddata.es/sparql	29884998	360	212	41	26(0), 12(100)
http://en.openei.org/sparql	27317782	1600	5163	306	209(0), 81(100)
http://datos.bcn.cl/sparql	52057935	542	357	122	157(0), 132(100)
http://id.eaufrance.fr/sparql	17743557	76	318	56	58(0), 46(100)
http://ldf.fi/warsa/sparql	14385118	90	310	57	110(0), 96(100)
http://ldf.fi/yoma/sparql	6627922	267	123	34	60(0), 53(100)

[1] L.Lāce et.al. Visual Presentation and Summarization of Linked Data Schemas. In: Proc. of KGSWC 2024, LNCS, Vol.15459

Algorithms for fragment computation

- Direct (paths from the root only)
- Heuristic (at each moment add the next most relevant node, may be order-sensitive wrt. tie breaking)
- Personalized PageRank (computed over entire graph)
- Stepwise Personalized PageRank (personalized PageRank over fragment and candidate nodes in each iteration)

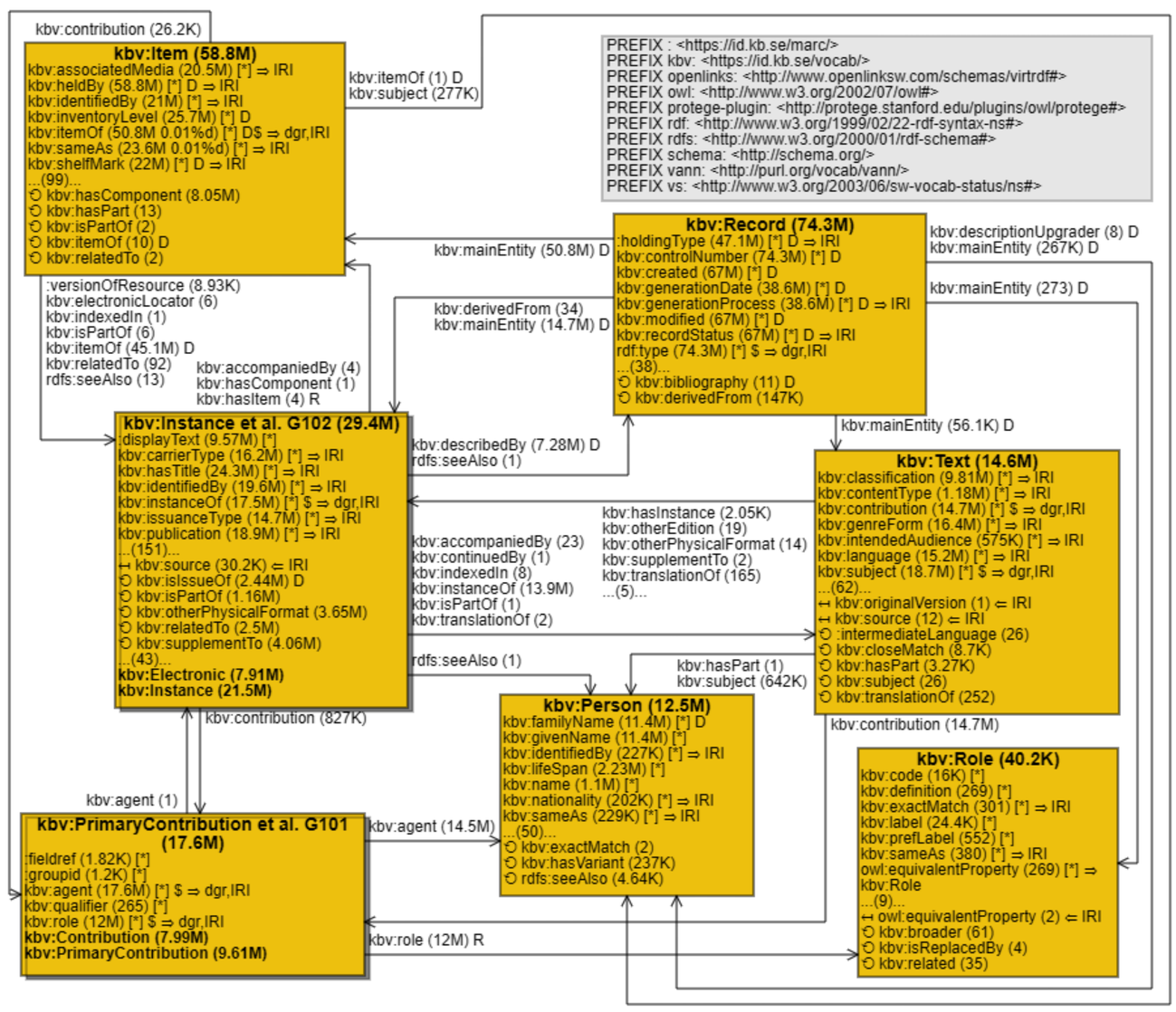
Weight computation by the edge relevance in the context of:

- Source class connections
- Source and target class connections
- Source and target class size

Further work:

Extra constraints, systematic evaluation

libris.kb.se: a fragment related with *kbv:Person* class



Two fragment extraction pipelines

- For medium-sized schemas, available in ***Viziquer*** schema diagram environment: ***within the diagram configuration tool.***
  Integrated in ViziQuer.
  https://github.com/LUMII-Syslab/viziquer-tools
- For large schemas (e.g., DBPedia): compute the fragment ***directly from the data set*** (SPARQL Endpoint) then import the fragment into the visual tool as a ***separate visual environment***.
  May be time consuming.
  https://github.com/LUMII-Syslab/kg-fragment-extractor

Visual schema and query resources: https://viziquer.lumii.lv/