

On the Analysis of Large Integrated Knowledge Graphs for Economics, Banking, and Finance

Under submission at EcoFinKG 2022

Shuai Wang

Xmas Hacking Project 1/2
KR&R Group, Vrije Universiteit Amsterdam

17th January, 2022

Content

1. Introduction
2. Integration
3. Analysis
4. Discussion



- Most ontologies are developed for domain-specific tasks.
- We have more data that requires inter-disciplinary knowledge.
- Economics, Finance, Banking are three closely related domains.
- We study the following aspects of the integrated knowledge graph:
 - 1 statistical analysis, alignment, and integration
 - 2 identity links
 - 3 (pseudo-)transitive relations



Related Work

Introduction

Integration

Analysis

Discussion

- 2007-10 US subprime mortgage crisis / 2007-08 global financial crisis
- the impact of the COVID-19 outbreak on Italian companies (Bellomarini et al.)
- customer banking (Shao et al.)
- company ownership
- public health policy
- social resilience



Introduction

Integration

Analysis

Discussion

- **Finance:** Financial Industry Business Ontology (FIBO), Financial Regulation Ontology (FRO), Insurance Regulation Ontology (IRO), Fund Ontology (FO), etc.
- **Economics:** STW (Standard Thesaurus Wirtschaft) Thesaurus, JEL classification system, etc.
- **Banking:** The WBG Taxonomy, Bank Regulation Ontology (BRO), etc. **Many are not open source projects.**



The KGs (10 graphs + 1 mapping)

- 1 the Financial Industry Business Ontology (we collected the FIBO ontology using OWL and FIBO vocabulary using SKOS)
- 2 the Financial Regulation Ontology (FRO)
- 3 the Hedge Fund Regulation (HFR) ontology
- 4 the Legal Knowledge Interchange Format (LKIF) ontology
- 5 the Bank Regulation Ontology (BRO)
- 6 the Financial Instrument Global Identifier (FIGI)
- 7 the STW Thesaurus for Economics (and its mappings)
- 8 the Journal of Economic Literature (JEL) classification system
- 9 the Fund Ontology



Shuai: Hi Frank, I don't know how to perform ontology alignment. Should I include also this part?

Frank: That's the only weakness of this paper. So, just do it!



LogMap (web interface) without reasoner.

Table 1: Alignment of knowledge graphs

	FIBO- vD	FIBO- OWL	LKIF	FIGI	STW	JEL	Fund
FIBO-vD	-	599	1	147	12	204	11
FIBO-OWL	-	-	24	516	5	57	70
LKIF	-	-	-	1	0	0	23
FIGI	-	-	-	-	0	34	2
STW	-	-	-	-	-	2	0
JEL	-	-	-	-	-	-	1
Fund	-	-	-	-	-	-	-

Thanks Majid for your help!



The Integrated Graph

Table 2: General statistics of knowledge graphs

Name	V	E	Size
FIBO-vD	17,547	28,128	1.1MB
FIBO-OWL	103,208	249,901	3.5MB
FRO	104,643	325,688	4.2MB
HFR	14,235	34,771	508KB
LKIF	1,032	2,378	72KB
BRO	259,067	837,994	10.0MB
FIGI	12,180	16,434	201KB
STW	51,128	113,276	1.0MB
JEL	12,109	17,758	448KB
Fund	9,758	34,889	409KB
STW-mappings	78,399	177,603	1.4MB
alignment	1,698	2,327	47KB
integrated	620,718	1,820,262	23.0MB

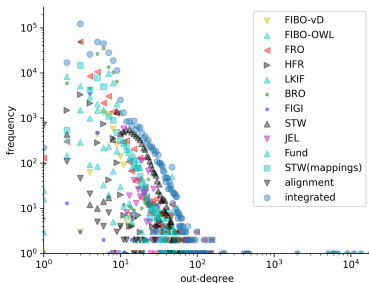
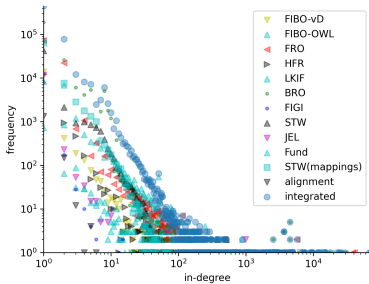


Table 3: Graph-theoretical statistics of knowledge graphs

Name	maxSCC	p_S (%)	maxWCC	p_W (%)
FIBO-vD	1	0.01	17,535	99.93
FIBO-OWL	297	0.29	103,208	100
FRO	17	0.02	104,443	99.81
HFR	849	5.96	14,230	99.96
LKIF	65	6.30	982	95.16
BRO	13	0.01	258,978	99.97
FIGI	13	0.11	12,180	100
STW	6777	13.25	51,128	100
JEL	1	0.01	12,099	99.92
Fund	411	4.21	9,750	99.92
STW-mappings	617	0.79	78,399	100
alignment	3	0.13	119	5.11
integrated	40,263	6.49	620,594	99.98



Analysis: Degree



Identity Links: CCs

Introduction

Integration

Analysis

Discussion

- 5,253 triples about owl:sameAs
- 29,556 triples about skos:exactMatch
- 8,172 triples about skos:relatedMatch
- 29,556 triples about skos:exactMatch
- 6,418 triples about skos:closeMatch



Identity Links: biggest WCCs

- owl:sameAs: 8 and 6
- skos:exactMatch: 119 and 45
- skos:relatedMatch: 21
- skos:closeMatch: 52

Not big at all. Can be manually refined.



Compare against LOD-a-lot

Introduction

Integration

Analysis

Discussion

- LOD-a-lot has much more owl:sameAs than skos:broaderMatch.
- LOD-a-lot has much more nested cycles for (pseudo-)transitive relations.
- LOD-a-lot needs (semi-)automatic tools for the refinement.
- They show different graph-theoretical features.



Thank You for your attention!

This project received help from Majid, Xu, Jan-Christoph, and Frank!

Contact: shuai.wang@vu.nl

