



A Knowledge Graph for Assessing Aggressive Tax Planning Strategies

Niklas Lüdemann¹, Ageda Shiba¹, Nikolaos Thymianis¹, Nicolas Heist¹ , Christopher Ludwig^{2,3} , and Heiko Paulheim¹

¹ Data and Web Science Group, University of Mannheim, Mannheim, Germany
`{nluedema, agshiba, nthymian}@mail.uni-mannheim.de,`
`{nico, heiko}@informatik.uni-mannheim.de`

² Area Accounting and Taxation, University of Mannheim, Mannheim, Germany
³ Corporate Taxation and Public Finance Department, Leibniz Centre for European Economic Research, Mannheim, Germany
`christopher.ludwig@zew.de`

Abstract. The taxation of multi-national companies is a complex field, since it is influenced by the legislation of several states. Laws in different states may have unforeseen interaction effects, which can be exploited by allowing multinational companies to minimize taxes, a concept known as *tax planning*. In this paper, we present a knowledge graph of multinational companies and their relationships, comprising almost 1.5 M business entities. We show that commonly known tax planning strategies can be formulated as subgraph queries to that graph, which allows for identifying companies using certain strategies. Moreover, we demonstrate that we can identify anomalies in the graph which hint at potential tax planning strategies, and we show how to enhance those analyses by incorporating information from Wikidata using federated queries.

Keywords: International taxation · Tax haven · Tax planning · Knowledge graph · Graph anomaly · Federated query

1 Introduction

Multinational corporations (MNCs), such as Google, IKEA, and Apple, have been scrutinized in the recent decade for so-called “aggressive” tax planning strategies. Taxes have a considerable effect on the net income of corporations, and it is in principle in the best interest of MNCs to reduce their worldwide tax burden by relocating profits within their group to lower-taxed affiliates.

The increasing internationalization of business activities in combination with the growing importance of the digital economy can create conflicts for the taxation of business profits by local governments [21]. For cross-border businesses’ activities, an appropriate allocation of foreign and domestic profits – and the underlying capital – to the involved jurisdictions is necessary, in accordance with the principle of economic allegiance [13]. MNCs represent an economic entity,

but they are usually organized as a conglomerate of legally independent separate legal entities or permanent establishments. The direct method to allocate profits and costs follows the separate entity approach and requires corporate divisions to behave as independent market participants, whereas the indirect method follows the unitary entity approach and allocates profits to affiliates by a formulary apportionment. The prevailing method in the international tax system is both for separate legal entities and permanent establishments the direct method which requires the application of the arm's length principle to intra-group transactions [20]. However, for many intermediary goods, services, and license contracts within MNCs, no independent reference market is observable and the implementation of the arm's length principle can be difficult.

Intuitively, MNCs have an incentive to allocate profits and costs in a tax-efficient way to reduce the overall tax burden of the corporation [7]. Tax reduction has a positive effect on the consolidated net income of MNCs which increases shareholder value. Efficient tax systems are – in theory – required to be neutral regarding any investment decision, but the diverse application of international taxation principles leads to a considerable heterogeneity between national tax systems [2]. Taxes represent costs for corporations, thus, MNCs usually consider tax effects intensively and pursue substantive and formal tax planning activities to change and structure economic activities in a tax-efficient way.

The term *tax planning* refers to generally accepted strategies to minimize tax liabilities of MNCs. Up to now, it is not precisely defined which tax planning strategies are considered as “aggressive”. The Organization for Economic Co-operation and Development (OECD) defines them as planning activities with “unintended and unexpected tax revenue consequences” [19]. In general, “aggressive” tax planning strategies are said to be in line with legal provisions but these strategies might be able to considerably reduce the tax burden of MNCs in some regions. In the following, the term “aggressive” refers to legal tax planning strategies of MNCs that lead to a substantial reduction of their tax liabilities [11]. Tax planning has to be differentiated from the terminology of *tax avoidance* and *tax evasion*. Tax avoidance strategies exploit loopholes in the tax law to reduce the tax liability. Tax evasion refers to any illegal activities to minimize the tax burden (e.g. misstatements in the tax declaration) [7].

MNCs are usually not one business entity, but a network of parent and child companies and holdings across different countries. Therefore, they can be directly represented in a knowledge graph (KG) [5], i.e., a graph describing entities and their relations [23]. In such a KG, companies can be connected among each other as well as to the countries they belong to, and further information (such as companies' legal forms, countries' populations and GDP etc.) can be added. Such a KG allows for two kinds of analyses: First, companies using certain aggressive tax planning strategies can be identified in the graph, since they correspond to characteristic subgraph patterns. Second, the graph can be analyzed for anomalies, which might hint at tax avoidance strategies, which are not yet known.

The rest of this paper is structured as follows. Section 2 describes the knowledge graph used for our analysis and its sources. Section 3 demonstrates the

above mentioned use cases, i.e., the identification of aggressive tax planning strategies and the search for graph anomalies. Section 4 discusses relevant related work, and Sect. 5 closes with a summary and an outlook on future work.

2 Knowledge Graph

For our analysis, we combine data from different sources into a knowledge graph, which can then be queried for analytics purposes.

2.1 Data Sources

The main source of our KG is the Global Legal Entity Identifier Foundation¹. GLEIF collects data from different legal entity identifier (LEI) issuers and provides a consolidated collection of that data. For each legal entity, different data fields (such as address, legal form, etc.) are collected. GLEIF has two levels of data: level 1 data (*who is who*) contains data about the companies as such, whereas level 2 data (*who owns whom*) provides information about the relationships between companies.

The level 2 data contains both direct as well as ultimate subsidiaries, i.e., child companies of child companies and so on. The latter is, in theory, equivalent to following the transitive closure of the subsidiary relation, however, in some cases, there are subsidiaries missing in between in the data for various reasons (e.g., country specific regulations for disclosing that information).

For further analyses, we include economic and geographic data for the entities at hand. To that end, country-specific data from the World Bank² and Wikidata [31] is collected. Those country-wide indicators include population and GDP. Moreover, we included the statutory corporate tax rate for each country from the OECD corporate tax database³.

Since some data was imported from Wikidata, we also provide interlinks to Wikidata. Countries and companies were trivial to match, since for the former, the GLEIF dataset uses ISO codes also present in Wikidata⁴, whereas for the latter, GLEIF identifiers are also used in Wikidata.⁵ Using that approach, we could interlink all countries and a total of 20,734 companies to Wikidata.

For matching cities, first, candidates are retrieved from Wikidata based on postal codes. To that end, a list of all entities with postal codes was retrieved from Wikidata, and attribute values with ranges are preprocessed to get an actual map of postal codes to entities (e.g., Berlin has only one value for the postal code attribute with value 10115-14199⁶). To deal with entities that do not represent a city (e.g., streets or libraries) and with cases where multiple candidates exist

¹ <https://www.gleif.org/en/>.

² <https://data.worldbank.org/>.

³ <https://www.oecd.org/tax/tax-policy/corporate-tax-statistics-database.htm>.

⁴ <https://www.wikidata.org/wiki/Property:P297>.

⁵ <https://www.wikidata.org/wiki/Property:P1278>.

⁶ <https://www.wikidata.org/wiki/Q64>.

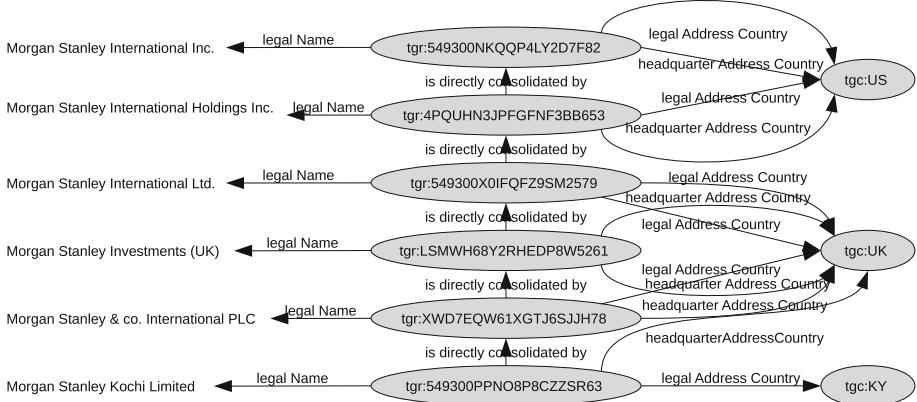


Fig. 1. Example representation of a company and its direct parents in the taxation graph

(e.g., 1000 is the postal code for Brussels, Sofia, Ljubljana, among others), the matching was made based on edit distance, with a maximum threshold of 0.3. Using that approach, we were able to link 43,832 cities to Wikidata.⁷

One basic design decision is collecting the data in one knowledge graph, vs. using SPARQL federated queries for Wikidata and Worldbank data. After some initial experiments with Virtuoso's query federation functionality, we found that federated queries are possible, but significantly slower than local queries. Hence, we follow a mixed approach: data about central entities (such as the population and GDP for countries) are included in our knowledge graph, while still maintaining the possibility to use the full data in Wikidata via federation.

2.2 Resulting Graph

The resulting graph contains about 1.5 M companies and 180k relationships between those companies, as shown in Table 1. An example representation of a company is shown in Fig. 1. Overall, the graph has 22,839,123 triples and is stored in a Virtuoso RDF store [8]. The knowledge graph is available online for browsing, download, and querying via a SPARQL endpoint.⁸

As depicted in Fig. 3, the distribution of direct and ultimate children follows a power law distribution. There are a few companies with very high number of ultimate children, as shown in Table 2, whereas the majority has only one or no ultimate children, as shown in Fig. 11. Companies with children have on average 2.6 direct children and 4.1 ultimate children (i.e., members of the transitive closure of the child relation). The longest chains of subsidiaries that we find

⁷ The full code for generating the knowledge graph is available online at <https://github.com/tax-graph/taxgraph>.

⁸ <http://taxgraph.informatik.uni-mannheim.de/>.

Table 1. Contents of the Knowledge Graph

<i>Class</i>	<i>Count</i>
Company	1,491,143
Country	225
City	95,306
Legal Form	1,286
<i>Relation</i>	<i>Count</i>
Direct subsidiary	87,020
Ultimate subsidiary	96,465

Table 2. Top 10 Companies with the most ultimate children

Company	No. of ultimate children
The Goldman Sachs Group, Inc.	2,534
Deutsche Bank Aktiengesellschaft	885
Morgan Stanley	793
Citigroup Inc.	686
Lloyds Banking Group PLC	680
Aegon N.V	629
The Royal Bank of Scotland Group Public Limited Company	496
HSBC Holdings PLC	472
Siemens Aktiengesellschaft	455
Societe Generale	429

spans across six companies, as shown in Fig. 1: Here, the ultimate child has a legal address in the Cayman Islands.

Figure 4 shows the distribution of legal and headquarter addresses. While the distribution among the top legal and headquarter addresses is similar, we can observe that two tax havens, i.e., Cayman Islands (KY) and British Virgin Islands (VG), appear among the top legal addresses, but not among the top headquarter addresses. For 36,400 of all companies in the graph (2.4%), the headquarter and legal address country differ; the majority of legal addresses in this set are the Cayman Islands (9,838), British Virgin Islands (5,878), Ireland (2,496), and Luxembourg (2,389). The most common combination is a headquarter address in the USA and a legal address in the Cayman Islands, as depicted in Fig. 2.

When comparing the corporate tax rates in the legal and headquarter addresses' countries, it can be observed that the corporate tax rate in the legal address country is, on average, 0.24 % points lower than in the headquarter's country. When considering only the 36,400 companies with differing addresses, that difference is even 10.5 % points. As depicted in Fig. 6, companies having their headquarter and legal address in different countries have a higher tendency of using a legal address in a lower-tax country.

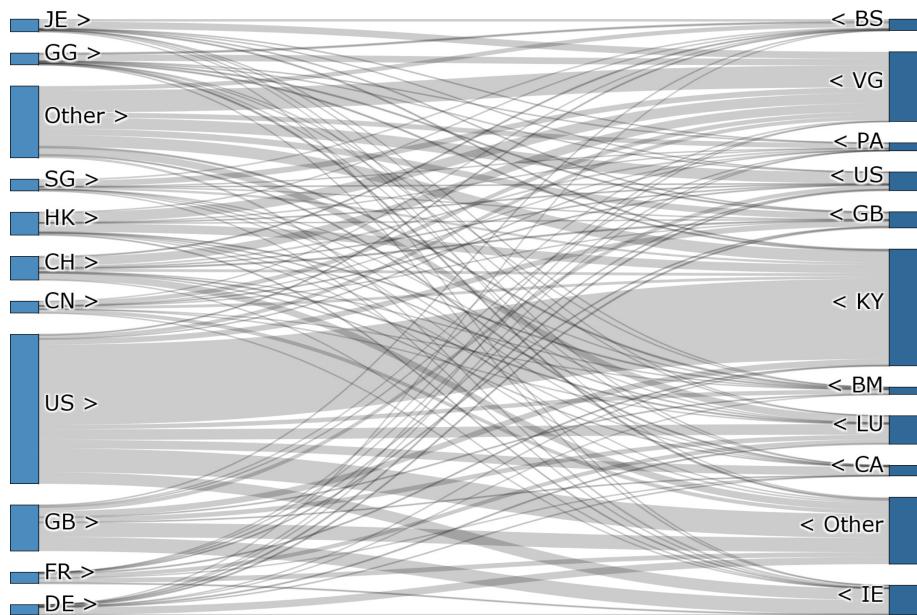


Fig. 2. Most frequent headquarter (left) and legal address country (right) for companies where headquarter and legal address are in different countries.

For subsidiary relations between companies, 35.7% of those are multinational, i.e., the legal address country of the subsidiary and its affiliate differ. Figure 5 depicts the most common relations for such multinational relationships. It can be observed that Ireland, India, and Singapore appear among the top 10 subsidiaries, but not among the top 10 parents.

When looking at the corporate tax rates for multinational companies, it can again be observed that the tax rate in which the subsidiary is located is typically lower than the one of the consolidating company. Across all subsidiary relations, the corporate tax rate in the child company's country is by 0.62 % points lower than in the parent company's country; if restricting this to multinational relations (i.e., where the parent and child company have their legal address in different countries), the difference is 2.46 % points.

3 Usage Examples

The knowledge graph can be used both for finding evidence for well-known tax avoidance strategies, as well as for searching for anomalies in the graph which hint at avoidance strategies not yet known.

3.1 Tax Avoidance Strategies

Well-known strategies for tax avoidance can be observed in the graph and formulated as query patterns and graph queries.

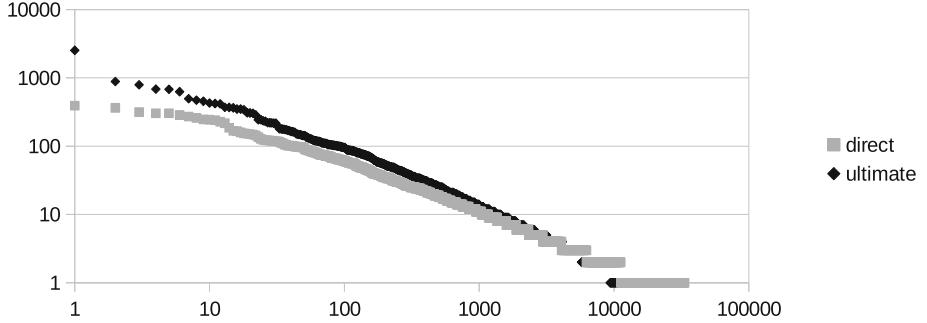


Fig. 3. Distribution of the number of direct and ultimate children per company

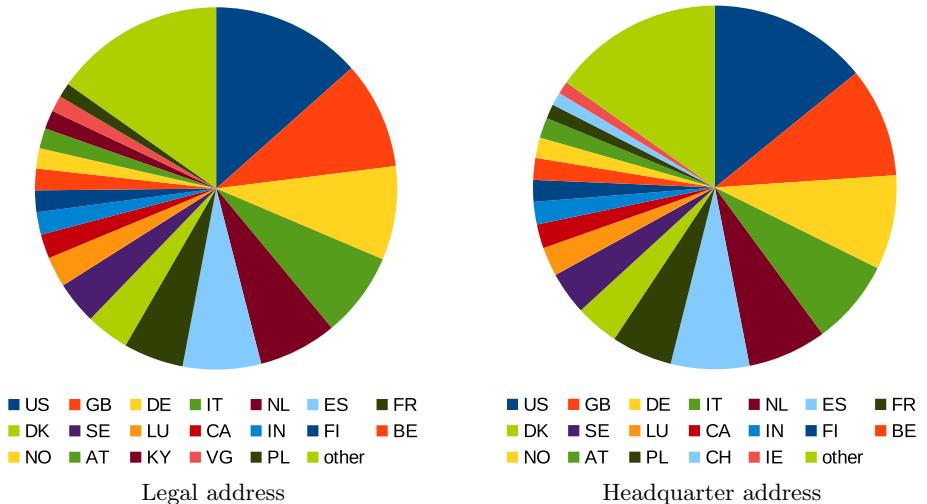


Fig. 4. Distribution of legal and headquarter addresses

Double Irish with a Dutch Arrangement. The Double Irish with a Dutch Arrangement uses in essence three companies: Two companies are located in Ireland (company A and C) and a conduit entity in the Netherlands (company B). Yet, the Irish fiscal authority considers only company A as taxable in Ireland, the second company is tax resident in a tax haven (company C). This allows to attribute all revenues to a tax haven (company C) [10,14].

Figure 7 depicts the query for a Double Irish with a Dutch Arrangement. Note that since further intermediate companies might be involved, we allow for chains of ownership by using `tgp:isDirectlyConsolidatedBy+`. Since the data in our knowledge graph is not complete, we could not find direct evidence for the Double Irish with a Dutch Arrangement construct. However, removing the last condition of the query (i.e., that company C has to have its headquarter in

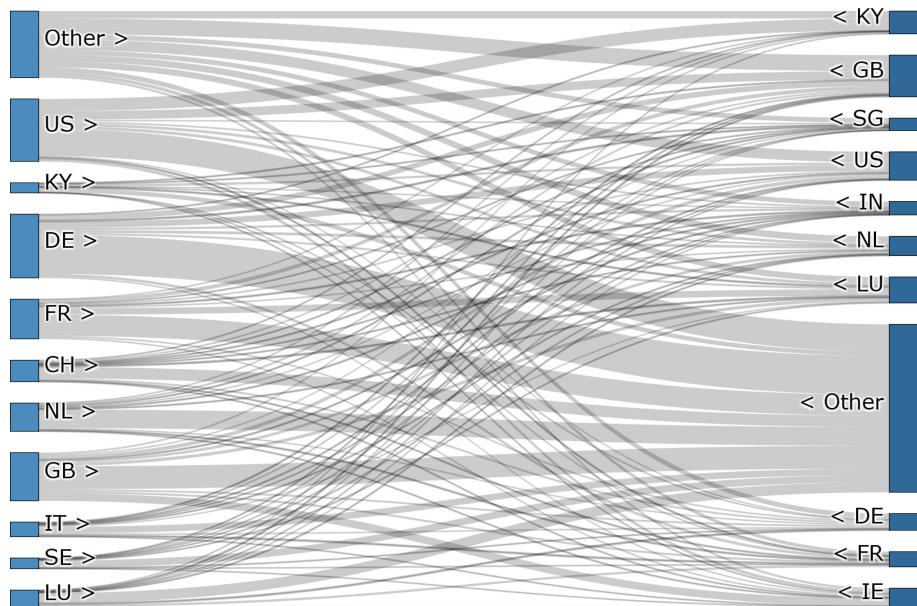


Fig. 5. Most frequent headquarter of parent (left) and child company's legal address country (right) for multinational subsidiary relations.

Ireland) yields 19 results (with the headquarter of C being located in countries such as the UK, the US, Japan, or Finland), which might hint at other variants of that tax planning strategy.

Duck-Rabbit Construct. Countries implement different legislative regulations which can have the unintended consequence that hybrid entities emerge. The OECD considers hybrid entities as firms with a dual residency and no country recognizes the entity as taxable [16, 22]. These constructs are called *duck-rabbit construct* in the following, named after the optical illusion in which some people see a duck, and some see a rabbit⁹. The structure can be as follows: a company C in the Netherlands having the legal form of a BV (a private limited partnership) is the child of a company B in a tax haven, which in turn is the ultimate child of some international company A, usually located in the US. In that case, the Dutch laws consider B a company under US tax legislation, while the US laws consider B a company under Dutch tax legislation, which ultimately leads to the company being taxed in none of the two countries.

The corresponding graph pattern and query are shown in Fig. 8. Running this query against the graph returns three constructs using the Bermudas and one using the Cayman Islands as an offshore tax haven. Among the former, there

⁹ https://en.wikipedia.org/wiki/Rabbit-duck_illusion.

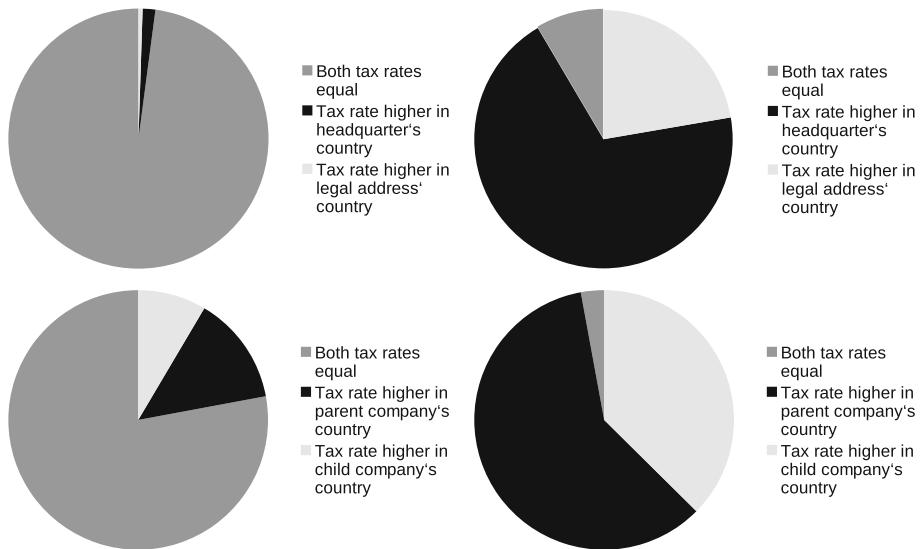


Fig. 6. Corporate tax differences between headquarter and legal address (upper part) and parent and child companies' addresses (lower part). The left hand side diagrams depict all companies, the right hand side diagrams are filtered to those where the two countries are different.

is also the game company *Activision*, which has become one of the well-known examples for this kind of tax avoidance strategy.¹⁰

3.2 Graph Anomalies

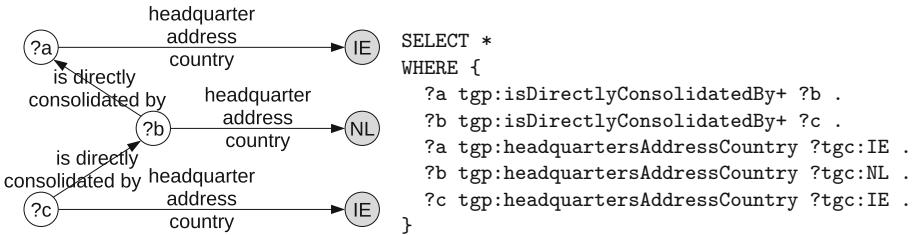
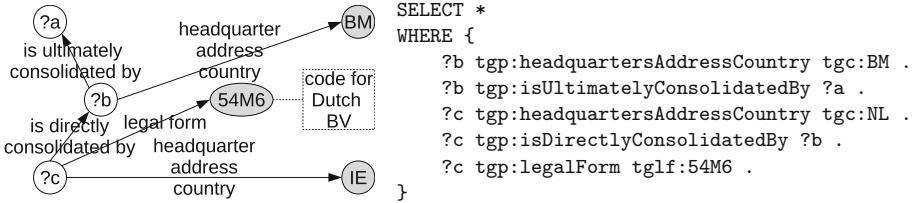
Since we included additional data about countries in our graph, we can use this as background information for further interesting observations [27]. One of those observations is the density of companies per state.

Table 3 depicts the top 10 countries by companies per capita and companies per GDP. It can be observed that many known tax havens appear in the top positions, with some values being clearly out of range (e.g., Liechtenstein lists one company per three inhabitants).

In the table of companies per capita, Denmark appears to be a bit of an outlier at first glance. Digging a bit deeper, we found that private holding companies – so called *Anpartselskab* – in Denmark are not taxed under certain conditions, and the creation of such companies is even advertised as a means for tax planning.¹¹ While this finding was new to the domain experts in the team, and we have not been able to fully explain the Denmark anomaly, we can, as of today, only find that “something is rotten in the state of Denmark” [29].

¹⁰ <https://thecorrespondent.com/6942>.

¹¹ See, e.g., <https://www.offshorecompany.com/company/denmark-holding/>.

**Fig. 7.** Double Irish Arrangement**Fig. 8.** Duck rabbit construction

Another analysis we conducted is related to addresses with high concentrations of companies using that address as a legal address. There are quite a few addresses which are used as legal addresses by thousands of companies. Examples for such addresses are shown in Fig. 9.

A particular observation of this analysis is that the two addresses most frequently used as legal addresses are in the state of Delaware, USA. We found that 36.7% of all US companies in our knowledge graph have their legal address in Delaware, whereas the state only accounts for 0.29% of the total US population. This phenomenon became known as the *Delaware Loophole* [32] and is a result of the Delaware tax legislation, which does not charge income tax on companies not *operating* in Delaware [4]. Consequently, only 15.3% of the companies having their legal address in Delaware also have their *headquarter* in that state.

3.3 Federated Querying

Although, as discussed above, federated queries for combining data from our knowledge graph with data from Wikidata are not very fast and scalable, they are still possible. One example is to use the area of cities – which is included in Wikidata but not in our KG – and compute the density of companies by headquarter and legal address in each city. The rationale is that cities exposing an overly large density are suspicious, similar to the analysis of addresses above.

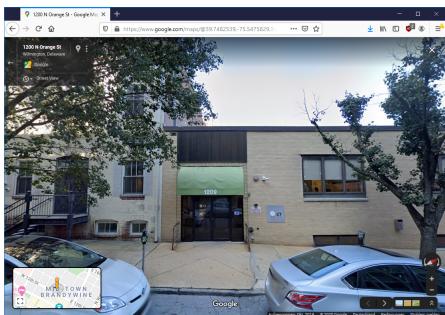
Figure 10 depicts an example for a federated query using Wikidata. The inner query collects all cities with a minimum number of companies using that city in their address, the outer query retrieves the area for those cities from Wikidata to compute the density of companies in those cities. Table 4 shows the outcome

Table 3. Top 10 countries by companies per inhabitants (top) and per GDP (in Million USD, bottom). Germany and USA are listed for comparison.

Country	Population	Companies per capita
Liechtenstein	37,910	0.311
Cayman Islands	64,174	0.237
Luxembourg	607,728	0.063
Isle of Man	84,077	0.036
Bermuda	63,968	0.035
Monaco	38,682	0.017
Marshall Islands	58,413	0.017
Seychelles	96,762	0.011
Denmark	5,797,446	0.009
Saint Kitts and Nevis	52,441	0.008
<i>Germany</i>	82,927,922	0.002
<i>USA</i>	327,167,434	0.001
Country	GDP	Companies per 1M GDP
Marshall Islands	221,278	4.59
Cayman Islands	5,1413	2.96
Liechtenstein	6,214	1.90
Seychelles	1,590	0.66
Belize	1,871	0.61
Samoa	820	0.57
Saint Vincent and the Grenadines	811	0.54
Luxembourg	70,885	0.54
Saint Kitts and Nevis	1,011	0.46
Isle of Man	6,770	0.45
<i>Germany</i>	3,947,620	0.03
<i>USA</i>	20,544,343	0.01

of that query, showing the top 10 cities according to the density of headquarter and legal addresses registered. It can be observed that in both cases, Vaduz in Liechtenstein has the highest density of companies per square kilometer. For the density of legal addresses, Dover in Delaware shows up in the top list as another piece of evidence for the already mentioned *Delaware Loophole*.¹²

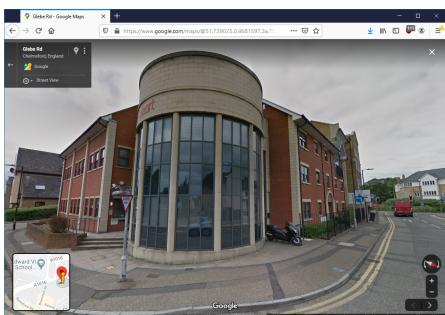
¹² The top 10 lists, however, have to be taken with a grain of salt. For a city to appear in the top 10 list, it requires that (a) we are able to link it to Wikidata using the approach sketched in Sect. 2, and (b) it has to have its area as a value in Wikidata. Therefore, those lists cannot be considered complete.



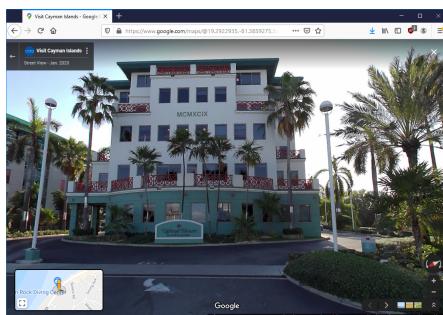
1209 Orange Street, Wilmington, Delaware, USA (14,551 companies)



251 Little Falls Drive, Wilmington, Delaware, USA (11,207 companies)



Eastwood House, Glebe Road, Chelmsford, CM1 1QW, Essex, United Kingdom (2,265 companies)



121 South Church Street, George Town, Cayman Islands (1,639 companies)

Fig. 9. Addresses with the highest frequency of being used as a legal address. Pictures from Google Street View.

4 Related Work

Parts of GLEIF, which we also used in this paper, have already been ported to an RDF representation and made available as a Linked Data endpoint [30]. However, the most important information for our use case – i.e., parent and child relations between companies – are not included in that representation.

Other approaches are restricted to single branches and/or countries, and thus would not allow for an analysis like the one conducted in this paper. An example for a branch specific solution is discussed in [9], where the authors build a populated ontology of bank holding companies and their ownership relations is introduced. The authors build an ontology and populated it from the Federal Reserve's public National Information Center (NIC) database¹³. Examples for country-specific solutions include a knowledge graph of Chinese companies [17], and a Linked Data endpoint of French business register data [6]. The euBusi-

¹³ <https://www.ffiec.gov/NPW>.

```

SELECT ?c ?count ?a (?count/?a as ?density) WHERE {
  { SELECT COUNT(?x) AS ?count ?c WHERE {
    ?x tgp:headquartersAddressCityID ?c .
  }
  GROUP BY ?c
  HAVING(COUNT(?x)>1000)
}
?c owl:sameAs ?wdc .
SERVICE <https://query.wikidata.org/bigdata/namespace/wdq/sparql> {
?wdc <http://www.wikidata.org/prop/direct/P2046> ?a}
} ORDER BY DESC(?density)

```

Fig. 10. Example for a federated query using Wikidata

nessGraph [28] project publishes data about businesses in the EU, but does not contain relationships between companies. Those datasets are often very detailed, but are of limited use for analyzing the taxation of multinational companies.

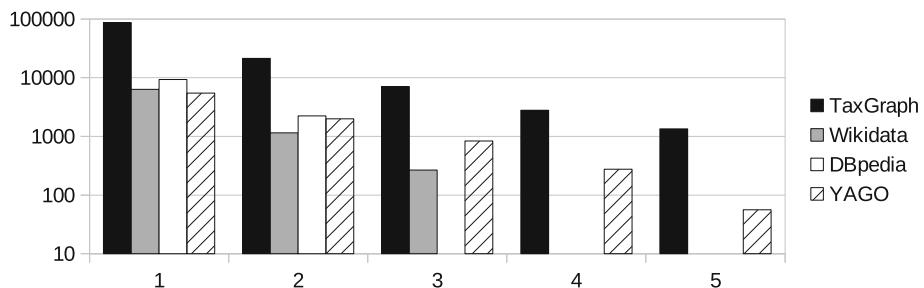
In addition to specific datasets, many cross-domain knowledge graphs also contain information about companies [26]. Hence, we also looked at such knowledge graphs as potential sources for the analysis at hand. However, since we need information not only for the main business entities, but also for smaller subsidiaries in order to identify tax compliance issues, we found that the information contained in those knowledge graphs is not sufficient for the task at hand. In Wikidata [31], DBpedia [15], and YAGO [18], the information about subsidiaries is at least one order of magnitude less frequent than in the graph discussed in this paper, as shown in Fig. 11: Especially longer chains of subsidiary relations, which are needed in our approach, are hardly contained in public cross-domain knowledge graphs.

In the tax accounting literature several scholars have already used data on multinational corporations to analyze the behavior of firms. It has been shown that some firms fail to publicly disclose subsidiaries that are located in tax havens [3]. In [1], the authors used very detailed data on the structure of multinational corporations to show that the introduction of public country by country reporting – the requirement to provide accounting information for each country a firm operates in to tax authorities – leads to a reduction in tax haven engagement.

A different strand of the tax literature has looked at networks of double tax treaties. Double tax treaties are in general bilateral agreements between countries that lower cross-border taxes in case of international transactions of multinational corporations. This literature on networks shows that some countries are strategically good choices for conduit entities to relocate profits and minimize cross-border taxation [12, 25].

Table 4. Cities with largest densities of companies having registered their headquarter (upper half) and legal address (lower half)

City	Country	No. of companies	Area in sq. km	Density
Vaduz	Liechtenstein	9021	17.30	521.45
Puteaux	France	1334	3.19	418.18
Paris	France	16276	105.40	154.42
Geneva	Switzerland	2153	15.92	135.24
Brussels	Belgium	3756	33.00	113.82
Copenhagen	Denmark	7356	86.70	84.84
Barcelona	Spain	8305	101.30	81.98
Milan	Italy	12563	181.67	69.15
Zug	Switzerland	1027	21.61	47.52
Nicosia	Cyprus	2276	51.06	44.58
Vaduz	Liechtenstein	8460	17.30	489.02
Puteaux	France	1218	3.19	381.82
Dover	USA (Delaware)	11268	60.82	185.28
Paris	France	16253	105.40	154.20
Brussels	Belgium	4213	33.00	127.67
Geneva	Switzerland	1522	15.92	95.60
Copenhagen	Denmark	7432	86.70	85.72
Barcelona	Spain	8095	101.30	79.91
Milan	Italy	12599	181.67	69.35
Zug	Switzerland	1052	21.61	48.68

**Fig. 11.** Chains of subsidiaries by number of hops

5 Conclusion and Outlook

In this paper, we have introduced a knowledge graph for multinational companies and their interrelations. We have shown that the graph allows for finding companies using specific constructs, such as well-known aggressive tax planning strategies, as well as for identifying further anomalies.

Our current knowledge graph uses company data from GLEIF, which is openly available and encompasses about 1.5 M business entities. There are other (non-open) databases such as ORBIS [24], which contain even more than 40M business entities, but their licenses do not allow for making them available as a public knowledge graph. For the future, we envision the dual development of an open and a closed version of the graph, the latter based on larger, but non-public data. On the larger graph, we expect to find more evidence for known tax planning strategies and a larger number of interesting anomalies.

Another interesting source of information would be the mining of up to date information from news sites such as Reuters or Financial Times. This would allow feeding and updating the KG with recent information, and to directly rate events in the restructuring of multinational companies in the light of whether or not it is likely that those events happen for reasons of tax planning.

Apart from increasing the mere size of the graph, we also plan to include more diverse data in the graph. For example, adding branch information for companies would allow for more fine-grained analyses finding tax planning strategies, which are specific to particular branches. Further data about companies could include the size of companies (in terms of employees), or other quantitative revenue data mined from financial statements, and a detailed hierarchy of subsidiary relations describing the relations more closely (e.g., franchise, licensee, holding).

A particular challenge lies in the more detailed representation of taxation legislation. For the moment, we have only included average corporate tax rates as a first approximation, but having more fine grained representations in the knowledge graph would be a clear improvement. However, this requires some up-front design considerations, since the ontological representation of tax legislation is not straight forward.

References

1. De Simone, L., Olbert, M.: Real Effects of Private Country-by-Country Disclosure (ID 3398116) (2019). <https://papers.ssrn.com/abstract=3398116>
2. Devereux, M., Freeman, H.: A general neutral profits tax. *Fiscal Stud.* **12**(3), 1–15 (1991)
3. Dyring, S., Hoopes, J.L., Langetieg, P., Wilde, J.H.: Strategic Subsidiary Disclosure (ID 3137138) (2018). <https://doi.org/10.2139/ssrn.3137138>, <https://papers.ssrn.com/abstract=3137138>
4. Dyring, S.D., Lindsey, B.P., Thornock, J.R.: Exploring the role delaware plays as a domestic tax haven. *J. Financ. Econ.* **108**(3), 751–772 (2013)
5. Ehrlinger, L., Wöß, W.: Towards a definition of knowledge graphs. *SEMANTiCS (Posters, Demos, SuCCESS)* **48**, 1–4 (2016)
6. El Kader, S.A., et al.: Modeling and publishing French business register (sirene) data as linked data using the eubusiness graph ontology (2019)
7. Endres, D., Spengel, C.: International Company Taxation and Tax Planning. Kluwer Law International (2015)
8. Erling, O.: Virtuoso, a hybrid RDBMS/graph column store. *IEEE Data Eng. Bull.* **35**(1), 3–8 (2012)
9. Fan, L., Flood, M.D.: An ontology of ownership and control relations for bank holding companies. In: Proceedings of the Fourth International Workshop on Data Science for Macro-Modeling with Financial and Economic Datasets, pp. 1–6 (2018)

10. Fuest, C., Spengel, C., Finke, K., Heckemeyer, J.H., Nusser, H.: Profit shifting and “aggressive” tax planning by multinational firms: issues and options for reform. *World Tax J.* **5**(3), 307–324 (2013)
11. Heckemeyer, J.H., Spengel, C.: Maßnahmen gegen steuervermeidung: Steuerhinterziehung versus aggressive steuerplanung. *Wirtschaftsdienst : Zeitschrift für Wirtschaftspolitik* **93**(6), 363–366 (2013)
12. Hong, S.: Tax Treaties and Foreign Direct Investment: A Network Approach. Working Paper (2016)
13. Jacobs, O.H., Endres, D., Spengel, C., Oestreicher, A., Schumacher, A.: Internationale Unternehmensbesteuerung: deutsche Investitionen im Ausland; ausländische Investitionen im Inland. Beck-Online Bücher, C.H. Beck, 8, neu bearbeitete und erweiterte auflage edn. (2016). https://beck-online.beck.de/?vpath=bibdata/komm/JacobsHdbIntStR_8/cont/JacobsHdbIntStR.htm
14. Kleinbard, E.D.: Stateless income. *Florida Tax Rev.* **11**, 699–773 (2011)
15. Lehmann, J., et al.: Dbpedia-a large-scale, multilingual knowledge base extracted from wikipedia. *Semant. Web* **6**(2), 167–195 (2015)
16. Lüdicke, J.: “tax arbitrage” with hybrid entities: challenges and responses. *Bull. Int. Taxation* **68**(6), 309–317 (2014)
17. Ma, Y., Crook, P.A., Sarikaya, R., Fosler-Lussier, E.: Knowledge graph inference for spoken dialog systems. In: 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5346–5350. IEEE (2015)
18. Mahdisoltani, F., Biega, J., Suchanek, F.M.: Yago3: a knowledge base from multilingual wikipedias (2013)
19. OECD: Study into the Role of Tax Intermediaries. OECD Publishing, Paris (2008)
20. OECD: Model tax convention on income and on capital (2014)
21. OECD: Addressing the Tax Challenges of the Digital Economy, Action 1–2015 Final Report. OECD/G20 Base Erosion and Profit Shifting Project, OECD Publishing (2015). <http://dx.doi.org/10.1787/9789264241046-en>
22. OECD: Neutralising the Effects of Hybrid Mismatch Arrangements, Action 2–2015 Final Report. OECD Publishing (2015)
23. Paulheim, H.: Knowledge graph refinement: a survey of approaches and evaluation methods. *Semant. Web* **8**(3), 489–508 (2017)
24. Ribeiro, S.P., Menghinello, S., De Backer, K.: The OECD orbis database: Responding to the need for firm-level micro-data in the OECD. *OECD Statistics Working Papers* **2010**(1), 1 (2010)
25. van’t Riet, M., Lejour, A.: Profitable Detours: Network Analysis of Tax Treaty Shopping (2015)
26. Ringler, D., Paulheim, H.: One knowledge graph to rule them all? Analyzing the differences between DBpedia, YAGO, Wikidata & co. In: Kern-Isbner, G., Fürnkranz, J., Thimm, M. (eds.) KI 2017. LNCS (LNAI), vol. 10505, pp. 366–372. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-67190-1_33
27. Ristoski, P., Paulheim, H.: Analyzing statistics with background knowledge from linked open data. In: SemStats@ ISWC (2013)
28. Roman, D., Soylu, A.: Enabling the European business knowledge graph for innovative data-driven products and services. *ERCIM News* **121**, 31–32 (2020)
29. Shakespeare, W.: Hamlet. Clarendon Press, Oxford (1912)
30. Trypuž, R., Kuzinski, D., Sopek, M.: General legal entity identifier ontology. In: JOWO@ FOIS (2016)
31. Vrandečić, D., Krötzsch, M.: Wikidata: a free collaborative knowledgebase. *Commun. ACM* **57**(10), 78–85 (2014)
32. Wayne, L.: How delaware thrives as a corporate tax haven. In: New York Times, vol. 30 (2012)