Tracing Organisation Evolution in Wikidata

Marieke van Erp, Jiaqi Zhu, and Vera Provatorova
DHLab KNAW Humanities Cluster
Amsterdam, the Netherlands
{marieke.van.erp,jiaqi.zhu}@dh.huc.knaw.nl

Abstract

Entities change over time, and while information about entity change is contained in knowledge graphs (KGs), it is often not stated explicitly. This makes KGs less useful for investigating entities over time, or downstream tasks such as historical entity linking. In this paper, we present an approach and experiments that make explicit entity change in Wikidata. Our contributions are a mapping between an existing change ontology and Wikidata properties to identify types of change, and a dataset of entities with explicit evolution information and analytics on this dataset.

1 Introduction

Already in 500BC Greek philosopher Heraclitus said that everything is in motion and nothing stays fixed (Graham, 2007). Knowledge Graphs (KGs) aim to capture information about entities and relationships between them, often modelling information from an entity-centric perspective (Rospocher et al., 2016). However, information about entity change is often not stated explicitly in KGs (Runge and May, 2023). This makes them less useful for investigating entity change over time or for downstream tasks such as entity linking to historical sources (Agarwal et al., 2018; Zaporojets et al., 2022). While there has been much work on entity evolution across different KGs (cf. (Halpin et al., 2010), we are looking at representations of change within a single KG.

To more usefully represent the evolution of entities, we have argued for 'unflattening' knowledge graphs (Van Erp, 2023). We define unflattening as representing the different aspects of an entity and how it changed over time, i.e. its evolution. In this paper, we present an approach and experiments to make explicit the evolution of organisation entities in Wikidata (Vrandečić and Krötzsch, 2014).[1]

[1] https://wikidata.org

We do so by expanding on properties that indicate some type of change, such as an acquisition or a change in the legal structure of an organisation. We thus make explicit the evolution of an entity using existing information in the knowledge graph.

Our main contributions are: 1) a mapping between Wikidata properties and a change ontology, 2) a set of organisation entities that exhibit change, and 3) analysis of the extracted entities.

The remainder of this paper is organised as follows. In Section 2, we discuss related work, followed by our framework and method in Section 3. The evolution of entities that we extract from Wikidata is described in Section 4, followed by a mapping of change types of Wikidata properties in Section 5. We present our analysis of the data in Section 6 and discussion in Section 7. We conclude with a summary and directions for future work in Section 8. Our code and data are available at `https://github.com/trifecta-project/wikidata-change`.

2 Related Work

Our work relates to modelling change of entities represented in KGs. In this section, we discuss the different research perspectives: data models for KGs that capture temporal information, evolution of entities across KGs, and modelling changes on top of an existing dataset.

Various ontologies have been proposed to model change, cf. (Welty and Fikes, 2006; Kauppinen and Hyvönen, 2007; Giménez-García et al., 2017). These ontologies often propose to add temporal information to a triple or set of triples to timebound a statement or set of statements. (Rospocher et al., 2016) proposed to model knowledge graphs from an event-centric instead of an entity-centric perspective, thus putting change at the forefront. The CIDOC-CRM (Doerr, 2005) is entity-centric but provides many building blocks to model things that happened to these entities.

Modeling the evolution of entities between different iterations of knowledge graphs (KGs) has been investigated in (Tasnim et al., 2019) and (Zhang et al., 2022). To go beyond a pairwise comparison of entities in two different snapshots of a KG, (Tasnim et al., 2019) automatically generate summaries of entities over different snapshots and compare these via a matrix. They focus on person entities. In (Zhang et al., 2022) the evolution of temporal knowledge graphs is investigated through a representation learning framework that takes into account both changing relations between entities as well changes in the overall structure of the KG. In our work, we focus on the changes in entities that are expressed within the same version of the KG, i.e. any temporally bound properties that express a change, rather than the evolution of the KG. This is termed the 'Temporal KG' or 'Time as data' perspective in Polleres et al. (2023).

Modelling change has been researched in the geographical information science (GIS) domain (cf. Stapel, 2023; Myrda et al., 2020; Bernard et al., 2018). Due to the long history of maps and digitisation efforts, GIS practitioners have been mapping changes in the scope of territories (e.g. the expansion and contraction of the Prussian empire) and names (e.g. Constantinople vs Istanbul) for decades. Changes can be recorded and connected in a pointwise manner but not qualified, as for example the Historical Atlas of the Low Countries 1350-1800 does (Stapel, 2023). Here, the focus is on creating and linking layers by points where each point has certain properties and points can be grouped into various clusters that designate an administrative area at any given time. Similarly, (Myrda et al., 2020) developed a conceptual schema for connecting different manifestations of a settlement over time that also includes properties to express name changes. Conceptually the closest model, and the one we adapt for expressing entity change more generally, is the TSN-Change ontology (Bernard et al., 2018) which is Linked Data-ready (Bernard et al., 2022).

While there is certainly an aspect of concept drift or concept change (Bloomfield, 1983) associated with the problem we are investigating, that research avenue focuses on how users perceive or use a concept. The research gap we address is the manner in which a concept change is represented within a KG: How can we make this change explicit such that it can be analysed?

3 Modelling Change

We take inspiration from modelling change in the field of geographical information systems (GIS) where (Bernard et al., 2018, 2022) defined change drivers for geographical entity as coming from *structure changes* and *feature changes*. A structure change denotes a change that impacts several features at the same time. In their case, a feature is a territory, in our case it could be an organisation or another entity type. A feature change denotes changes that only affect one feature (i.e. territory), such as a name change. These changes can be mapped to two of the three aspects of concept drift as defined by (Wang et al., 2011). They define a label as how the entity or concept is referred to, its intension are the properties, or characteristics, implied by it, and the extension the set of things that are covered by a concept. A name change corresponds to a shift on the *label* of a concept, whilst structural changes correspond to changes of the definition or *intension* of a concept. We consider the use or *extension* of a concept out of the scope of this work as we focus on properties inside a KG. We have also considered ontologies specific to the organisation domain such as BORO (de Cesare and Partridge, 2016) and COOT (Bogea Gomes et al., 2023) that model organisations and their transformations. However, for the modelling of organisational structures, we take the Wikidata data model as a given and we prefer a more general description of changes rather than one very specific to the organisation domain to facilitate extensions to other domains in future work.

Not all feature changes have to be sequential, as entities can also hold different roles at the same time. As not all elements of the TSN Change ontology presented in (Bernard et al., 2018) apply to non-geographical entities, we focus on those classes that are relevant to other types.

StructureChange Change operations that impact the entity and several of its features simultaneously. An example of structure change in organisations is when one organisation acquires another one.

Merge Two or more entities merge, e.g. two political parties merge;

Split An entity splits into two or more entities e.g. a band splits up and the band members go on to individual music careers.

FeatureChange A change that affects one entity of a given type. An example is when an organisation changes its name or when a person starts a different role, for example from being a lawyer to a politician.

- **Appearance** The entity comes into existence.
- **Disappearance** The entity ceases to exist.
- **IdentificationChange** The manner in which the entity is identified changes.
  - **IdentifierChange** The entity's identifier changes, e.g. when bank identifier codes (BINs) change;
  - **NameChange** The entity's name changes, e.g. a football player changes his name;
  - **DescriptionChange** The entity's description changes e.g. a company's motto changes.
- **GeometryChange** The entity changes in size, shape, or structure.
  - **Expansion** The entity grows larger, e.g. a city acquires more land;
  - **Contraction** The entity becomes smaller, e.g. the number of employees a company has decreases.
  - **Deformation** The entity changes shape, e.g. a company branches out into different industries.

4 Extracting Change from Wikidata

In this paper, we focus on organisations and specifically their subtypes businesses, companies and enterprises. Figure 1 illustrates our workflow.

To identify a set of change indicators for businesses, we first query for all organisations and properties associated with them. This results in a set of 3,211 properties which are manually assessed for whether they indicate a change and if so, what type of change. Further details on this mapping are described in Section 5. For this, we used the Wikidata mirror hosted by University of Freiburg at `https://qlever.cs.uni-freiburg.de/wikidata/` as it has a longer time-out than the main Wikidata endpoint. There is a trade-off here as its information is not always up to date as we found by querying for number of unique entities of type organisation on both endpoints (347,357 on the Freiburg endpoint vs. 473,292 on the Wikidata endpoint).[2] However, we deem the set of results on the Freiburg endpoint large enough to assess general properties and statistics on this entity type and the Wikidata public endpoint resulted in a time-out. The results from this step were used to identify the most relevant change properties. All other queries from hereon are queried via the Wikidata public endpoint. We do take a subset of organisations, namely businesses, companies, and organisations for these, to manage the query load.

[2]Query performed on 7 March 2025.

5 Mapping Change Types to Properties

3,211 Unique Wikidata properties are present in our dataset. For the annotator, we also display the property label and description from the Wikidata ontology. Each of these is annotated with two labels: a change type, and whether the property inherently describes change. For the change types, we discern the following:

Appear	Disappear
Merge	Split
Expand	Contract
Identifier Change	Name Change
Description Change	Deformation
Not a change property	

The second label that is assigned, expresses whether a property describes a change, such as P571 inception *(time when an entity begins to exist)* or whether only multiple instances of that property at different points in time express a change, such as P169 chief executive officer *(highest-ranking corporate officer appointed as the CEO within an organization)*. The number of times a type of change is assigned to a property is shown in Table 2. Mapping the properties took one annotator about 5 hours.

The majority of the properties do not inherently express a change (3,160). Of those 51 that do, 22 indicate some type of appearance or coming into existence such as P577 publication date or P1619 date of official opening and 16 that express some disappearance or ending such as P570 date of death and P576 dissolved, abolished or demolished date.

Due to Wikidata's interconnectedness to other resources, it contains a large number of identifiers to other databases such as P1417 Encyclopædia Britannica Online ID and P2025 Find a Grave cemetery ID which, if a Wikidata resource is mapped to one identifier at one point in time, and another at another point in time, indicates an identifier change. Another large portion of the mappings express po-

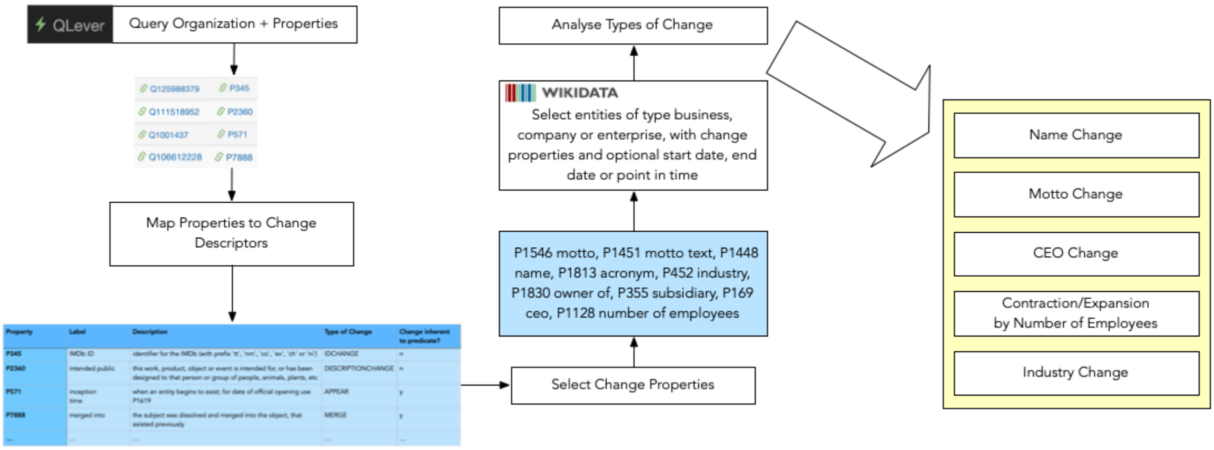

Figure 1: Overview of steps to extract and analyse change of business, company, and enterprise data from Wikidata

Type of information	# present	# empty	% present
Start time (P580)	11,204	228,001	4.684
End time (P582)	6,331	232,874	2.647
Point in time (P585)	15,217	223,988	6.361

Table 1: General statistics on the number of time-bound statements about businesses, companies and enterprises in Wikidata

tential *Description Changes* such as P286 head coach and P452 industry *(specific industry of company or organization)*.

Some properties express a type of change going either way such as expanding or contracting P2351 number of graves *(in a graveyard)*, or even more complex changes such as P1830 owner of which can express a merger, a split, an expansion or a contraction, for example in the case of a company acquiring or selling other companies.

6 Analysing Change

To characterise changes across business entities in Wikidata, we chose to focus on 6 different aspects of businesses that may change over time: 1) name/acronym, 2) motto, 3) industry, 4) ownership (of other organisations), 5) ceo, and 6) number of employees (indicating expansion or contraction).

For each of these dimensions, we collect entities involved and, if present, start and/or end times or points in time during which a statement was valid. To mitigate endpoint time-outs, we first query Wikidata for all entities of type business, then we iteratively perform a query for each of the 232,605 entities to obtain change indicators. All queries were performed in February and early March 2025 on the public Wikidata SPARQL endpoint. Not every entity has the properties that we query for, thus our resulting dataset contains 109k entities that have at least one statement that expresses a change in its name, motto, industry, ownership, ceo, or number of employees.

Temporal data in Wikidata is identified via qualifiers on statements. The most common temporal qualifiers in Wikidata are P585 for point in time, P580 for start time, and P582 for end time. Whilst other large, general purpose KGs such as DBpedia[3] or Yago[4] were also considered, an exploratory analysis showed that Wikidata contains the most time-bound information suited to our purposes such as corporate acquisitions accompanied by dates. This is probably due to the fact that Wikidata sources its information from a variety of sources. Yago 4.5 has incorporated parts of Wikidata to provide a cleaner and more consistent resource (Suchanek et al., 2024), we found that the type of information that expresses change in entities is less well represented in YAGO than in Wikidata. However, as Table 1 shows, most information regarding organisation in Wikidata does not have explicit timestamps associated with it. This is in line with earlier research on temporal information contained in Wikidata (Santos et al., 2024).

In the remainder of this section, we illustrate the

[3]https://dbpedia.org
[4]https://yago-knowledge.org/

Change Type	Count
Inherent Change Property	51
Not an Inherent Change Property	3,160
Appear	27
Disappear	20
Merge	1
Split	1
Expand	1
Contract	0
Identifier Change	2,254
Name Change	37
Description Change	657
Deformation	59
Not a Change Property	68
Appear/Disappear	2
Merge/Split	1
Expand/Contract	73
Identifier/Description Change	1
Name/Description Change	2
Merge/Split/Expand/ Contract	4
Merge/Split/Expand/Contract/ Description Change	2
Appear/Disappear/Merge/Split/ Expand/Contract/ID/Description Change	1

Table 2: Count of the number of occurrences of each type of change in the Wikidata property mapping

different types of change via use cases from Wikidata. As most entities can only appear or disappear once, we focus on the Identification Changes and Geometry Changes.

Identification Changes

Companies are identified by their name, acronym but also their mottos and CEOs. Some mottos, such as Nike's 'Just do it' have become such strong markers of the company that they can be considered an identifier (Court et al., 1997). Chief executives are often the face of a company and as they can control its product, its image and culture (Bloom and Rhodes, 2018) we consider them part of the identifier of a company.

Name Change

The history of a company's name change can be a brief history of the company. In our dataset, we found 41,904 entities with a name change, and 1,174 with an acronym change. Table 3 presents an example of a company that has changed it name 7 times: Baldwin Locomotive Works Wikidata lists

Company name	Period
M.W. Baldwin	1825-1839
Baldwin, Vale & Hufty	1839-1842
Baldwin & Whitney	1842-1845
M. W. Baldwin	1854- 1867
M. Baird & Company	1867-1873
Burnham, Parry, Williams & Company	1873-1890
Burnham, Williams & Company	1891-1909
Baldwin Locomotive Works	1909-

Table 3: Name changes of Baldwin Locomotive Works

the names and dates for which these names were valid and we see that for example at least once it changed its name back to a previous name (M. W. Baldwin). Names sometimes also hint at mergers, acquisitions or partnerships such as when M.W. Baldwin became Baldwin, Vale & Hufty, a partnership that only lasted a few years after which the partners went separate ways and other partnerships were formed (Brown, 1995). By diving into the background of these name changes, it shows how the company adapted itself to changes in markets and society and how that is reflected in its name.

Description Change: Motto

A company motto change could be a sign of a change in company values or brand identity, or a reflection of a change in company structure. Our dataset contains 24 companies for which we have information regarding their motto changes. The majority of the motto changes are not associated with a date (18). For the ones that are accompanied by dates, the information is not applied consistently, for example for Google there are three motto statements, once expressed through property P1546 motto, the other two are expressed through the motto text property PP1451. Two of the statements have an end date, one statement has a start date associated with it. Some other entities have start and end dates, or only point in time dates. For some companies, such as Twinkl, only two of its motto statements are accompanied by start times, the other 8 do not have temporal information associated with them.

Description Change: CEO

Since a company's leadership can be seen as a part of its identifier, we turn to the timeline of CEO changes to trace the evolution of a company. Fig-

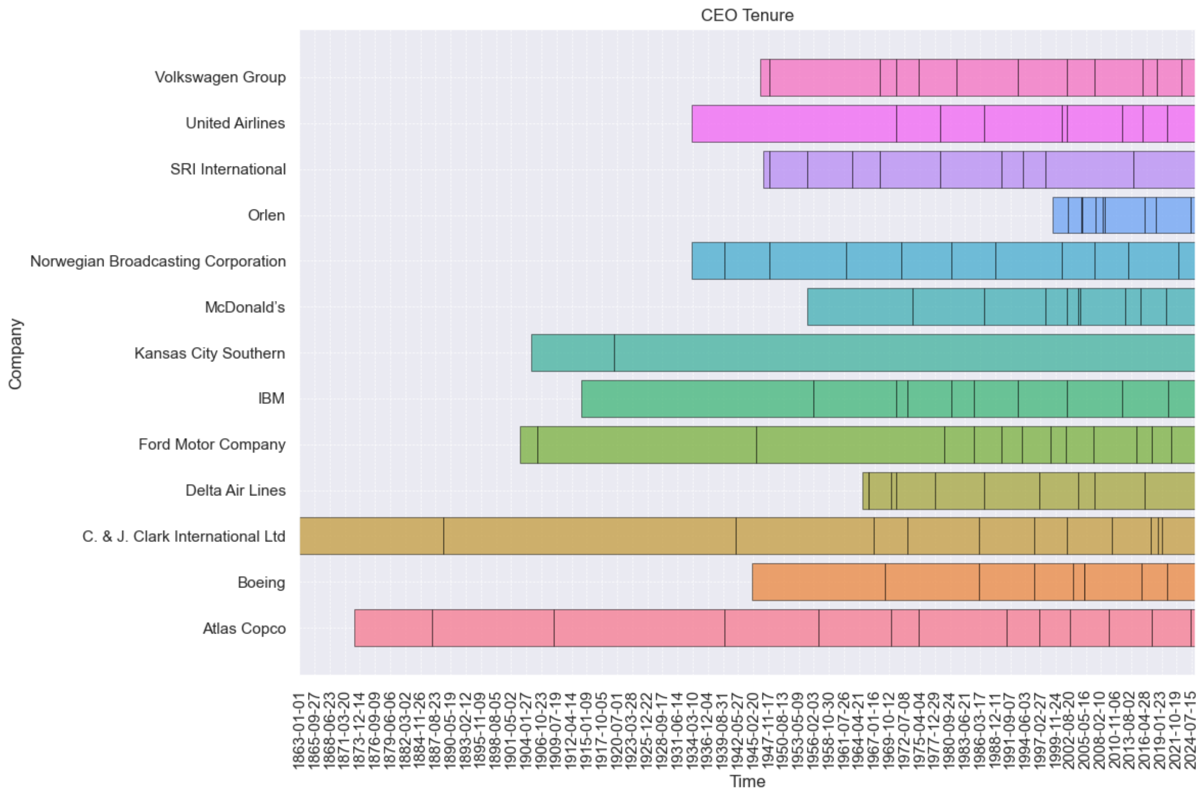

Figure 2: CEO tenure timeline in the companies with the most CEO changes.

ure 2 shows the CEO tenure timeline of companies that had the most CEO changes in our dataset. These companies were selected because their number of CEO changes was in the 95th percentile or higher. The figure indicates that the duration of CEO tenures in our dataset tends to decrease towards the end of the timeline. This suggests that the rate of identifier change for these companies is accelerating. One company where the contrast between long tenures of its early CEOs and more frequent later leadership changes is noticeable is C. & J. Clark International Ltd, more informally known as the Clarks shoe company. Clarks is shown in mustard yellow in Figure 2. This company started as a family business in 1863, the company was controlled by its first CEO for 26 years and its third CEO was in charge for 25 years. From the information available in Wikidata, it looks like the second CEO was in charge from 1889 until 1942, but during that time the company had multiple directors (all from the Clark family) alongside each other with distinct roles (Palmer, 2013). The Wikidata structure is not well equipped to handle a more-than-one-CEO-at-the-time structure, which highlights the complexity of capturing the real world in data and the difficulties in preserving data quality in KGs (Shenoy et al., 2022). After this, Clarks leadership changes became relatively more frequent, reaching a turbulent period in 2017 with two consecutive CEO tenures shorter than a year and stabilising again in 2019.

In contrast, a company deviating from the trend is SRI International, an American research institute (shown in lavender blue in Figure 2). Founded in 1946, the institute has undergone relatively frequent leadership changes at the beginning, with its longest CEO tenure spanning from 1998 to 2014.

Expansion/Contraction

Fluctuations in the number of employees of a business can be one indicator of a business expanding or contracting. To examine the potential expansion or contraction, we visualize the change in employee size over time for twenty companies with at least 25 data points for number of employees in Wikidata.

Figure 3 shows that the growth in number of employees for most businesses was relatively stable during 2000-2025. An outlier here is Deutsche Post AG, which shows quite some expansion as its line rises more steeply.

We calculated the regression line slope to get more than an eyeball estimate of a company's ex-

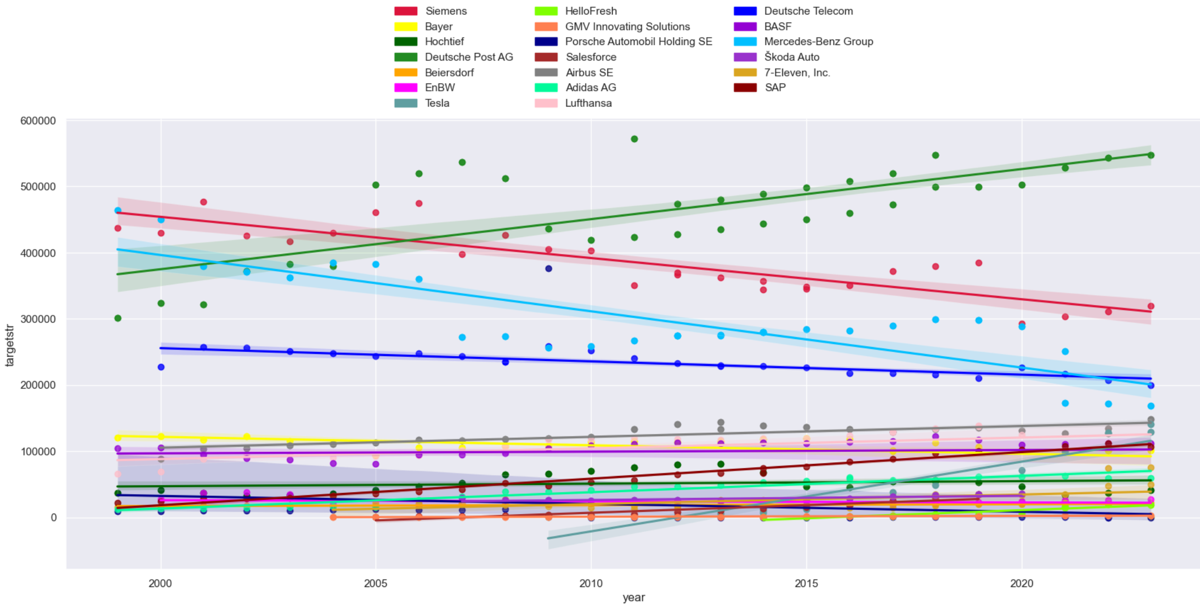

Figure 3: Change of employees over time during 2000-2024 for companies with most data points regarding employee numbers in Wikidata

pansion or contraction. In the case of a company's expansion/contraction, a positive slope indicates that the company expands during the given time period as its employee number rises, a negative slope indicates that the company experiences contraction during the given time period as its employee size shrinks. Greater slope values indicate greater expansion or contraction of the company.

In Table 4, the highest slope value we found is for Tesla, which can be explained by the increasing popularity of the electric vehicle industry.[5] An explanation can be found in the general increase in its number of employees between 2010 and 2024 is in line with its growing production, as Tesla works to provide electric vehicles for the broader market (Carlier, 17 April 2024). Another highly positive slope is found for Deutsche Post AG, a privatised mail and parcel shipment and delivery company based in Germany but with worldwide coverage. Since its privatisation in 1995 it has been steadily acquiring other businesses and expanding its operations. It has also benefited from the growth in online shopping (Thiele, 2024). The erratic behaviour of the employee numbers of Porsche Automobil Holding SE warrant further investigation. This company displays a negative slope value as well as a major outlier point (navy blue) in Figure 3 in 2009. When consulting the Porsche SE Annual Report 2022[6] which is provided as reference to the most recent employee number, it provides an insight into the complexity of the Porsche holding which in 2022 only had 38 employees, but as the owner of the Volkswagen Group it has many more employees within various company substructures. This is reflected in Wikidata as information about Porsche's subsidiaries and their employees is present. It is out of the scope of this paper to connect these, but it would present an interesting use case to do this.

Deformation

A company can be said to change its shape, or deform, when it branches out into a different industry or industries. There are various graph similarity measures that can be used to compute the distance between two nodes in a graph (cf. Rada et al., 1989; Caballero and Hogan, 2020). We use the topsim measure as defined in (Ilievski et al., 2024) to calculate the similarity between industries. The advantage of this measure is that it aggregates different similarity measures over the Wikidata graph.

As many companies have multiple industries associated with them. We compute the maximum, minimum and average distance between the main company's industries and its subsidiary industries. Table 5 displays companies for which we have at

[5] At the time of writing in March 2025, the Tesla company was decreasing in popularity, this data was not available yet in Wikidata.

[6] https://www.porsche-se.com/fileadmin/user_upload/PSE2022_Annual_Report_en.pdf Last visited: 21 March 2025

Entity	Slope
Mercedes-Benz Group	-23.284
Siemens	-17.039
Deutsche Telecom	-5.454
Bayer	-3.502
Porsche Automobil Holding SE	-3.281
enBW	-0.468
GMV Innovating Solutions	0.323
beiersdorf	0.376
BASF	0.689
Hochtief	1.053
Škoda Auto	1.549
7-Eleven, Inc.	4.036
Lufthansa	4.443
Airbus SE	4.516
Adidas AG	6.758
HelloFresh	6.614
SalesForce	6.322
SAP	11.025
Deutsche Post AG	20.711
Tesla, Inc.	29.497

Table 4: Calculation of the regression line slope to identify employee size growth. Slope values are given in ascending order, rounded up to the nearest thousandths

least 30 industry data points, meaning that the company and its subsidiaries have industries defined at least 30 times and have a low similarity between some of its initial business sector and its subsidiary business sector(s). This illustrates the complexity of the Wikidata structure and property assignments as for example *financial services* and *banking industry* have a much higher similarity score (0.908) *financial sector* and *banking industry* (0.557).

One of the companies that branches out most according to Wikidata is Google. In Table 6, the minimum distance between Google (industries: Internet, information technology, Internet marketing, software industry, web search engine, Internet industry) and its subsidiaries is shown. While the majority of the companies that Google acquired over the years are related to internet technology but there are some outliers such as aerial photography and robotics. It should be noted here that Wikidata listed other acquired companies, but not all had information regarding their industry available.

There are also companies that do not branch out into other industries through their subsidiaries, such as Van der Valk, a Dutch hospitality company, whose subsidiaries listed in Wikidata are hotels.

7 Discussion

The main limitation to our temporal information extraction experiments is the availability of data. The proportion of timebound statements in Wikidata is limited: (Santos et al., 2024) found that only 7.35% have a point in time associated with a statement, 4.96% have a start date, and 2.54% have an end time. In our experiments, we only used these, but Wikidata has an additional 64 other very specific dateTime-type properties, such as P9946: 'date of probate' or P574: 'year of publication of scientific name for taxon'. The YAGO consortium has already worked on consolidating the Wikidata data model (Suchanek et al., 2024), using a cleaned up data model would be preferable to creating query templates that cover the (current) 67 temporal properties. Event-centric databases such as Bio2RDF (Dumontier et al., 2014), NewsReader (Rospocher et al., 2016) and the European Olfactory Knowledge Graph (Lisena et al., 2022) provide more temporally bound information, but the trade-off is that they are more domain-specific and would thus be suited to more specific questions rather than general questions.

Furthermore, the coverage of entities is uneven. For example, for Finlayson, a Finnish textile manufacturer founded in 1825, the data regarding its number of employees has 20 statements but only up to 1925, whilst the company is still in business. The reason for this lies in the power of Wikidata being a collaborative KG, as this particular information comes from a Finnish history book published in 1932 that a user inserted,[7] but it does provide gaps and unpredictable coverage. When comparing Wikidata to other large-scale general KGs such as DBpedia and YAGO, it is considered the most complete for information regarding people, music albums and films, but less so for organisations, places, and events (Ringler and Paulheim, 2017).

The richness of the Wikidata data model and the size of the resource also provides a hurdle in collecting and analysing information. It has almost become imperative to run a local version of the resource to be able to query it effectively. The size of the Wikidata taxonomy and its inconsisten use mostly affects our industries analysis. This is a known problem that has yet to be resolved (Brasileiro et al., 2016).

[7] `https://www.wikidata.org/wiki/Q97898858`

Entity (industry)	Subsidiary (industry)	Topsim
Gazprom (gas)	Volzhsky synthetic fiber plant (light)	0.627
Unilever (food)	4P Rube Göttingen (packaging)	0.495
HSBC (financial)	HSBC Bank Taiwan (banking)	0.557
WarnerMedia (media)	Hanna-Barbera Studios Europe (animation)	0.396
Microsoft (software)	Mojang Studios (video games)	0.503
Van der Valk (horeca)	Theaterhotel Almelo (hotel)	0.551
Coca-Cola (consumer goods)	Odwalla (beverage)	0.576
Latvenergo (electricity retailing)	Ķeguma SES (Support services to forestry)	0.533
UniCredit (financial services)	Živnostenská banka (Other monetary intermediation)	0.655
General Electric (automotive)	NBC (communication)	0.365

Table 5: Sample of company, business, and enterprise entities with low similarity score to their subsidiaries' industries

Subsidiary name	topsim score	Industry
Kaltix Corp.	0.461	Internet
Jaiku	0.461	Internet
DoubleClick	0.437	Internet marketing
Fitbit	0.429	consumer electronics
Google Nest	0.405	home automation
Niantic	0.374	video game industry
Boston Dynamics	0.341	robotics

Table 6: Distance between Google and its subsidiaries according to their associated industry.

8 Conclusion & Future Work

In this paper, we have presented: 1) a mapping between over 3,000 Wikidata properties and an adapted version of the TSN Change ontology, 2) a set of organisation entities that exhibit change, and 3) an analysis of the extracted entities.

We show that the TSN Change ontology that was developed for the geographical domain can be applied more generally. We have extracted temporally bound information from Wikidata and classified it according to the change ontology, illustrating that whilst more information could be temporally bound, the available data already indicates that organisations exhibit different types of change and this is captured in the data. The entities we have extracted, along with over statements that describe some change of that entity provide a starting point for exploring how businesses, companies and enterprises change, which we have done in our analyses of companies and their subsidiaries, company names and acronyms, company mottos, leadership changes and number of employees over time.

In future work, we will extend our experiments to cover information about more different types of entities and change events that happen to them. As Wikidata has good coverage of people and music albums, those would be obvious domains to start. Furthermore, the historical domain provides additional entity types that exhibit change, such as ships that have changed roles and names. A 20th century example is MV Wilhelm Gustloff, a German cruise ship that was repurposed as a military ship or the 18th century French ship *Jason* which was captured by the English and sold to the Dutch who renamed it *Toevalligheid*.[8] People can undergo changes too such as different roles and/or titles, see for example Charles V. For this, additional information will need to be extracted from other sources and we intend to consult company histories and experts. We aim to feed this information back into Wikidata, along with enrichments about change descriptors such as the mappings after we have tested them on these additional entity types.

Richer data will also enable more in-depth analysis of triggers of change, for example how does a change in leadership correlate with other changes in the company such as acquisitions or number of employees. Across companies, one could trace whether companies 'copy' each other's behaviour or whether they maintain their own strategy.

Our ontology and experiments have opened up a new avenue of investigating temporal change in KGs showing that the 'Time as data' (Polleres

[8] https://resources.huygens.knaw.nl/das/detailVoyage/98119 Last visited: 21 March 2025.

et al., 2023) perspective is there, and there is much more to explore.

Author contributions (by author initials) are listed according to the Contributor Roles Taxonomy (CRediT). Conceptualization: ME; Data curation: ME, JZ; Formal Analysis: ME, JZ, VP; Funding acquisition: ME; Methodology: ME, JZ; Project administration: ME; Software: ME, JZ; Supervision: ME; Visualization: ME, JZ, VP; Writing (original draft): ME, VP; Writing (review and editing): ME, JZ.

Acknowledgments

Funded by the European Union under grant agreement 101088548 - TRIFECTA. Views and opinions expressed are however those of the author only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them. The authors would like to thank their colleagues Paul Groth, Jelle van Lottum, Manjusha Kuruppath, Arno Bosse, and Brecht Nijman for their helpful comments and suggestions.

References

Prabal Agarwal, Jannik Strötgen, Luciano del Corro, Johannes Hoffart, and Gerhard Weikum. 2018. diaNED: Time-aware named entity disambiguation for diachronic corpora. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 686–693, Melbourne, Australia. Association for Computational Linguistics.

Camille Bernard, Marlène Villanova-Oliver, Jérôme Gensel, and Hy Dao. 2018. Modeling changes in territorial partitions over time: ontologies tsn and tsn-change. In *Proceedings of the 33rd Annual ACM Symposium on Applied Computing*, SAC '18, page 866–875, New York, NY, USA. Association for Computing Machinery.

Camille Bernard, Marlène Villanova-Oliver, and Jérôme Gensel. 2022. Theseus: A framework for managing knowledge graphs about geographical divisions and their evolution. *Transactions in GIS*, 26(8):3202–3224.

Peter Bloom and Carl Rhodes. 2018. *CEO society: The corporate takeover of everyday life*. Bloomsbury Publishing.

Leonard Bloomfield. 1983. *An introduction to the study of language*. John Benjamins Publishing Company.

Silvia Bogea Gomes, Flavia Maria Santoro, Miguel Mira da Silva, Paulo Pinto, and Giancarlo Guizzardi. 2023. Towards a core ontology of organisational transformation. *Applied Ontology*, 18(1):31–70.

Freddy Brasileiro, João Paulo A Almeida, Victorio A Carvalho, and Giancarlo Guizzardi. 2016. Applying a multi-level modeling theory to assess taxonomic hierarchies in wikidata. In *Proceedings of the 25th international conference companion on World Wide Web*, pages 975–980.

John K. Brown. 1995. *The Baldwin Locomotive Works, 1831–1915: A Study in American Industrial Practice*. Studies in Industry and Society series. Johns Hopkins University Press, Baltimore, Maryland, USA.

Marco Caballero and Aidan Hogan. 2020. Global vertex similarity for large-scale knowledge graphs. In *Wikidata@ ISWC*.

Mathilde Carlier. 17 April 2024. Number of tesla employees from july 2010 to december 2023. https://www.statista.com/statistics/314768/number-of-tesla-employees/.

David C. Court, Anthony Freeling, Mark G. Leiter, and Andrew J. Parsons. 1997. If nike can "just do it," why can't we? *The McKinsey Quarterly*, 33(3):24–34. Copyright - Copyright McKinsey & Company, Inc. 1997; Last updated - 2023-11-29; SubjectsTermNotLitGenreText - United States–US.

Sergio de Cesare and Chris Partridge. 2016. Boro as a foundation to enterprise ontology. *Journal of Information Systems*, 30(2):83–112.

Martin Doerr. 2005. The cidoc crm, an ontological approach to schema heterogeneity. In *Semantic Interoperability and Integration*. Schloss-Dagstuhl-Leibniz Zentrum für Informatik.

Michel Dumontier, Alison Callahan, Jose Cruz-Toledo, Peter Ansell, Vincent Emonet, François Belleau, and Arnaud Droit. 2014. Bio2rdf release 3: a larger connected network of linked data for the life sciences. In *Proceedings of the 2014 international conference on posters & demonstrations track*, volume 1272, pages 401–404. Citeseer.

José M Giménez-García, Antoine Zimmermann, and Pierre Maret. 2017. Ndfluents: An ontology for annotated statements with inference preservation. In *The Semantic Web: 14th International Conference, ESWC 2017, Portorož, Slovenia, May 28–June 1, 2017, Proceedings, Part I 14*, pages 638–654. Springer.

Daniel W Graham. 2007. Heraclitus. In *Stanford Encyclopedia of Philosophy*. Stanford University - The Metaphysics Research Lab.

Harry Halpin, Patrick J Hayes, James P McCusker, Deborah L McGuinness, and Henry S Thompson. 2010. When owl: sameas isn't the same: An analysis of identity in linked data. In *The Semantic Web–ISWC 2010: 9th International Semantic Web Conference,* *ISWC 2010, Shanghai, China, November 7-11, 2010, Revised Selected Papers, Part I 9*, pages 305–320. Springer.

Filip Ilievski, Kartik Shenoy, Hans Chalupsky, Nicholas Klein, and Pedro Szekely. 2024. A study of concept similarity in wikidata. *Semantic Web*, 15(3):877–986.

Tomi Kauppinen and Eero Hyvönen. 2007. Modeling and reasoning about changes in ontology time series. In *Ontologies: A handbook of principles, concepts and applications in information systems*, pages 319–338. Springer.

Pasquale Lisena, Daniel Schwabe, Marieke van Erp, Raphaël Troncy, William Tullett, Inger Leemans, Lizzie Marx, and Sofia Colette Ehrich. 2022. Capturing the semantics of smell: the odeuropa data model for olfactory heritage information. In *European Semantic Web Conference*, pages 387–405. Springer.

Grzegorz Myrda, Bogumił Szady, and Agnieszka Ławrynowicz. 2020. Modeling and presenting incomplete and uncertain data on historical settlement units. *Transactions in GIS*, 24(2):355–370.

Mark Palmer. 2013. *Clarks: Made to Last: The story of Britain's best-known shoe firm*. Profile Books.

Axel Polleres, Romana Pernisch, Angela Bonifati, Daniele Dell'Aglio, Daniil Dobriy, Stefania Dumbrava, Lorena Etcheverry, Nicolas Ferranti, Katja Hose, Ernesto Jiménez-Ruiz, Matteo Lissandrini, Ansgar Scherp, Riccardo Tommasini, and Johannes Wachs. 2023. How Does Knowledge Evolve in Open Knowledge Graphs? *Transactions on Graph Data and Knowledge*, 1(1):11:1–11:59.

Roy Rada, Hafedh Mili, Ellen Bicknell, and Maria Blettner. 1989. Development and application of a metric on semantic nets. *IEEE transactions on systems, man, and cybernetics*, 19(1):17–30.

Daniel Ringler and Heiko Paulheim. 2017. One knowledge graph to rule them all? analyzing the differences between dbpedia, yago, wikidata & co. In *KI 2017: Advances in Artificial Intelligence: 40th Annual German Conference on AI, Dortmund, Germany, September 25–29, 2017, Proceedings 40*, pages 366–372. Springer.

Marco Rospocher, Marieke Van Erp, Piek Vossen, Antske Fokkens, Itziar Aldabe, German Rigau, Aitor Soroa, Thomas Ploeger, and Tessel Bogaard. 2016. Building event-centric knowledge graphs from news. *Journal of Web Semantics*, 37:132–151.

Lars Runge and Wolfgang May. 2023. Towards semantic identification of temporal data in rdf. In *Proceedings of the 34th GI-Workshop on Foundations of Databases (Grundlagen von Datenbanken)*, Hirsau, Germany. CEUR-WS.

Veronica Santos, Daniel Schwabe, and Sérgio Lifschitz. 2024. Can you trust wikidata? *Semantic Web*, Preprint:1–22.

Kartik Shenoy, Filip Ilievski, Daniel Garijo, Daniel Schwabe, and Pedro Szekely. 2022. A study of the quality of wikidata. *Journal of Web Semantics*, 72:100679.

Rombert Stapel. 2023. Historical atlas of the low countries: A gis dataset of locality-level boundaries (1350–1800). *Research Data Journal for the Humanities and Social Sciences*, 8(1):1–33.

Fabian M. Suchanek, Mehwish Alam, Thomas Bonald, Lihu Chen, Pierre-Henri Paris, and Jules Soria. 2024. Yago 4.5: A large and clean knowledge base with a rich taxonomy. In *Conference on Research and Development in Information Retrieval (SIGIR)*.

Mayesha Tasnim, Diego Collarana, Damien Graux, Fabrizio Orlandi, and Maria-Esther Vidal. 2019. Summarizing entity temporal evolution in knowledge graphs. In *Companion proceedings of the 2019 World Wide Web conference*, pages 961–965.

Sonja Thiele. 2024. Nachfragetrends im deutschen brief-und paketmarkt und deren auswirkungen auf den universaldienst. Technical report, WIK Diskussionsbeitrag.

Marieke Van Erp. 2023. Unflattening knowledge graphs. In *Proceedings of the 12th Knowledge Capture Conference 2023*, pages 223–224.

Denny Vrandečić and Markus Krötzsch. 2014. Wikidata: a free collaborative knowledgebase. *Commun. ACM*, 57(10):78–85.

Shenghui Wang, Stefan Schlobach, and Michel Klein. 2011. Concept drift and how to identify it. *Journal of Web Semantics*, 9(3):247–265.

Chris Welty and Richard Fikes. 2006. A reusable ontology for fluents in owl. In *Proceedings of the 2006 conference on Formal Ontology in Information Systems: Proceedings of the Fourth International Conference (FOIS 2006)*, pages 226–236.

Klim Zaporojets, Lucie-Aimée Kaffee, Johannes Deleu, Thomas Demeester, Chris Develder, and Isabelle Augenstein. 2022. Tempel: Linking dynamically evolving and newly emerging entities. *Advances in Neural Information Processing Systems*, 35:1850–1866.

Jiasheng Zhang, Shuang Liang, Yongpan Sheng, and Jie Shao. 2022. Temporal knowledge graph representation learning with local and global evolutions. *Knowledge-Based Systems*, 251:109234.