# OntoDrift: a Semantic Drift Gauge for Ontology Evolution Monitoring

Giuseppe Capobianco[1 [0000−0001−9702−8189]], Danilo Cavaliere[1 [0000−0003−2859−0447]], and Sabrina Senatore[1 [0000−0002−7127−4290]]

Department of Information and Electrical Engineering and Applied Mathematics, University of Salerno, Fisciano (SA) 84084, Italy
{dcavaliere,ssenatore}@unisa.it

**Abstract.** This paper presents OntoDrift, an approach to detect and assess the semantic drift among timely-distinct versions of an ontology. The semantic drift is evaluated at the concept level, by considering the main features involved in an ontology concept (e.g., intention, extension, labels, URIs, etc.) and at the structural level, by inspecting the taxonomic relations among concepts (e.g., subclass, superclass, equivalent class). New measures are defined to evaluate the semantic drift among individual concepts from different ontology versions, and among entire ontology versions. OntoDrift extends identity-based approaches to assess the drift among ontology versions not only on concepts in common among versions, but also on concepts added and removed during the ontology evolution to improve the drift assessment. OntoDrift can also be run over big-sized ontology versions, as shown in a case study about DBpedia. Experiences on various ontologies show the potential of OntoDrift in assessing the semantic drift among ontology versions.

**Keywords:** Semantic drift · Ontologies · similarity measures.



## 1 Introduction

An ontology allows the representation of knowledge on a domain of interest as a shareable, formal, and machine-understandable conceptualization. In many fields, such as video surveillance [2] and bioinformatics [1], where the knowledge domain tends to change over time, the ontology evolution process needs management. Since the ontology reflects the domain it describes, changes in the domain affect unavoidably the ontology dynamics. Changes in the domain imply changes to the meaning of concepts, which are generally referred to as semantic drifts [5, 8]. The changes affect the representation of concepts, as well as the relations among them across consecutive ontology versions. Automatic tools for semantic drift assessment are demanded to help experts in dealing with the tough, expensive and time-consuming ontology management. Semantic drift has been widely explored in linguistics [7], [4], but these methods focus on text instead of changes in the Semantic Web formalism. Some works [3], [8] explored the semantic drift among ontology versions by considering changes in both the structure and the content of the ontology. In [3], drift assessment is achieved by clustering ontology

population, other solutions [5] introduce linguistics-based methods to detect changes in the textual concept description, or exploit well-known model, such as the vector space model, to detect changes in concept features [10]. To assess the semantic drift among two ontology versions, two approaches are generally used: the morphing-chain and the identity-based [8]. The former compares each concept $A^i$, in the first ontology version $O_i$, to each concept $B^j$ in the second version $O_j$. The latter assumes that the concept identity is known, i.e., each concept $A^i$, in the ontology version $O_i$, is known to correspond (or match) to a unique concept $B^j$ in $O_j$. Both methods have advantages and drawbacks, in fact, the morphing-chain approach has very bad performances and is unsuited for big-sized ontologies [9]. The identity-based approach achieves better performances, but it does not consider unmatching concepts across versions in the drift assessment among two ontology versions [9]. Beyond the existing approaches, drift evaluation depends on the concept aspects considered. The morphing-chain framework in [8] assesses the drift among ontology versions on a concept including three aspects: label, intension and extension. This concept notion does not take into account many concept aspects, such as concept URI and taxonomy relations. Our approach, instead, introduces a new notion of concept, taking into account all these aspects; it extends the identity-based method with additional measures to provide a more refined assessment of the semantic drift at concept level and entire ontology version level.
The remainder of the paper is organized as follows: Section 2 focuses on the concept definition and the semantic drift measures. Section 3 is devoted to show the potential of the approach in the drift assessment. Section 4 highlights the benefits of the approach through comparisons with the reference framework. Last section is devoted to the conclusions.

## 2 Semantic Drift Assessment: notions and measures

OntoDrift has been designed to evaluate the semantic drift at concept and ontology levels. The approach defines the ontology *Concept* in terms of multiple aspects related to class name (e.g., labels), intensional and extensional aspects (i.e., properties and instances), *Concept* identity (e.g., URI) and structural relations, i.e. taxonomic relations with other concepts (e.g., equivalent classes, subclasses, superclasses, etc.). This notion of *Concept* involves many more distinct aspects, concerning both the *Concept* features and relations, compared with other approaches, that exclusively consider the label, the intensional and extensional aspects [8]. OntoDrift introduces new measures to assess the semantic drift over a concept considering its multiple aspects and between two ontology versions enahancing the identity-based approaches to consider the concepts added and removed throughout the ontology update.

### 2.1 The Concept and its Aspects

A concept is defined as an ontology class that can have properties and relationships with other concepts. A generic *Concept* is shown in Figure 1 along with its aspects used for assessing the ontology drift: the inherited ones from the reference approach [8] are in cyan, the extended ones are in yellow, while the new-defined aspects are

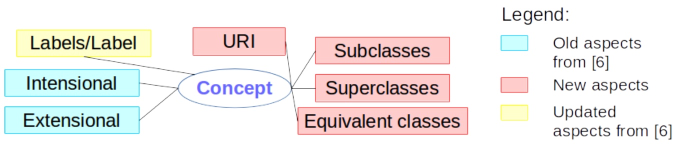


Fig. 1: A *Concept* schema with its aspects.

in red. Formally, let us suppose that $O^t$ is the ontology (version) updated at the date $t$; each concept is related to an object (another concept, a literal, etc.) according to the $\langle$*subject, predicate, object*$\rangle$ triple relation. Let us define the ontology version as $O^t = \langle C^t, R^t, I^t, T^t, V^t \rangle$, where $C^t$ is the set of classes or concepts, $R^t$ is the set of relations, $I^t$ is the set of individuals, $T^t$ is the set of data types, $V^t$ is the set of values; the aspects about the *Concept* $c^t \in C^t$ are defined as follows.

- **URI aspect**. A URI is a string that uniquely identifies a *Concept* $c^t \in C^t$. Formally:

$$c^t_{uri} = u \tag{1}$$

where $u$ is the subject (i.e., a URI) of the triple $\langle u, rdf{:}type, owl{:}Class \rangle_t$.
- **Labels aspect**. A set of the labels used to refer to a specific *Concept* $c^t$, also in different languages (when ontologies are multilingual). Each item of this set comes from all the objects contained in the *Concept* triples with the property $rdfs{:}label$. Let us define *labels* as:

$$c^t_{labl} = \{l | \langle c^t, rdfs{:}label, l \rangle_t)\} \tag{2}$$

where $l$ is text representing a label.
- **Intensional aspect**. A set of triples that have $rdfs{:}domain$ or $rdfs{:}range$ as predicate. Each triple links a property to the *Concept* through one of these two predicates. More formally, let $p$ be a generic property ($p \in R^t$), the intensional aspect of the *Concept* $c^t$ is defined as follows:

$$c^t_{ints} = \{c^t_d \cup c^t_r\} \tag{3}$$

where $c^t_d$ and $c^t_r$ are properties having $c^t$ as domain and range defined as follows:

$$c^t_d = \{p | \langle p, x, c^t \rangle_t, (x = rdfs{:}domain)\} \tag{4}$$

$$c^t_r = \{p | \langle p, x, c^t \rangle_t, (x = rdfs{:}range)\} \tag{5}$$

- **Subclasses aspect**. A set of URIs identifying *Concepts* that are explicit subclasses of a specific *Concept* $c^t$. It is created by taking the subject of the triples with the property *rdfs:subClassOf* as predicate and the analyzed *Concept* $c^t$ as object. Let us formally define the aspect as follows:

$$c^t_{sub} = \{s | \langle s, rdfs{:}subClassOf, c^t \rangle_t\} \tag{6}$$

where $s \in C^t$ is a triple subject (i.e., a class identified by a URI).

- **Superclasses aspect**. A set of URIs identifying ancestor *Concepts* of the analyzed *Concept* $c^t$ . The set is composed of the parent *Concepts* of the given *Concept* $c^t$. Let us define this aspect as follows:

$$c^t_{sup} = \{s|\langle c^t, rdfs{:}subClassOf, s\rangle_t\} \tag{7}$$

where $s \in C^t$ is a triple object representing a URI.
- **Equivalent classes aspect**. A set of URIs identifying all the *Concepts* equivalent to the *Concept* $c^t$, viz., all the objects in the triples, whose predicate is the property $owl{:}equivalentClass$ associated with $c^t$.

$$c^t_{eq} = \{e|\langle c^t, owl{:}equivalentClass, e\rangle_t\} \tag{8}$$

where $e \in C^t$ is a class (concept) identified by a URI.
- **Extensional aspect**. A set of URIs identifying all the individuals of the *Concept* $c^t$. Each individual is the subject of a triple linked to $c^t$ by the property *rdf:type*.

$$c^t_{ext} = \{x|\langle x, rdf{:}type, c^t\rangle_t\} \tag{9}$$

where $x \in I^t$ is a URI identifying an individual.

According to the concept aspects defined, a concept $c^t \in C^t$ of the ontology version $O^t$, can be described as follows:

$$c^t = \langle c^t_{uri}, c^t_{labl}, c^t_{ints}, c^t_{sub}, c^t_{sup}, c^t_{eq}, c^t_{ext}\rangle \tag{10}$$

### 2.2 Semantic drift assessment at concept level

The semantic drift among ontology versions is assessed by considering the drift on *Concept* pairs, where concepts in a pair belong to distinct ontology versions. OntoDrift introduces some similarity measures to assess the drift among two concepts. Given two *Concepts* $A$ and $B$, belonging to two ontology versions $O^t$ and $O^{t\prime}$ ($A \in C^t$ and $B \in C^{t\prime}$), the similarity measure on each of the aspects (introduced in Section 2.1), is defined as follows.

- **Similarity on the URI aspect**. The similarity on the URI aspect among two *Concepts* consists of checking whether or not the two *Concepts* have the same identifier, i.e., they describe the same resource. Recall that each URI in an ontology uniquely identifies a resource, that can be a *Concept*, a relation or an individual, a datatype, etc. Let us assume that if the concepts from different ontology versions have the same URI, they are identical. For this reason, the similarity on the URI aspect is 1 when the URIs coincide, otherwise the result is 0. Let $A$ and $B$ be two *Concepts*, the similarity on the URI aspect is defined as follows:

$$sim_{uri}\left(A_{uri}, B_{uri}\right) = \begin{cases} 1, & \text{if } A_{uri} = B_{uri} \\ 0, & \text{otherwise} \end{cases} \tag{11}$$

where $A_{uri}$ and $B_{uri}$ represent the URI aspects of the *Concept* $A$ and $B$, respectively.

– **Similarity on aspects labels, subclass, superclass, equivalent class and extensional**. The aspects are name sets, and are described by the Jaccard index [6], which evaluates the drift by counting how many instances (names) the two concepts have in common in relation to all their instances for an aspect. For each aspect in Equations (2), (6)-(9), the measure considers the set of elements (precisely, the element names) that describe that aspect. For example, if the aspect is the *label* (Equation (2)), the set of label names, associated with two concepts A and B, are compared. Similar evaluations can be applied on the other aspects: in general, considering the aspect $a$, among the possible aspect names: $\{labl, sub, sup, eq, ext\}$, the similarity value can be defined as follows:

$$sim_a\left(A_a, B_a\right) = \frac{|A_a \cap B_a|}{|A_a \cup B_a|} \tag{12}$$

where $A_a$ and $B_a$ are the name sets of the concepts $A$ and $B$ respectively, on the aspect whose name is $a$. The $sim_a$ values lie in the range $[0, 1]$, where 0 means no similarity among the two sets, and 1 represents the equality among the two sets (same set of names). The higher the value, the more the *Concepts* $A$ and $B$ are similar on the aspect $a$.

– **Similarity on the intensional aspect**. Since the intensional aspect involves triples whose predicate is one of $rdfs{:}domain$ and $rdfs{:}range$, the concepts are compared on the set of the domain or range instances, respectively. If $A$ and $B$ play the role of range in the triple $\langle p, rdfs{:}domain, c\rangle$ (i.e., $c = A$ or $c = B$) the similarity $sim_d$ is evaluated on the set of the domain properties for the concept $c$ (see Equation 4) by using the Jaccard index (Equation 12). Similarly, the similarity $sim_r$ between the two *Concepts* $A$ and $B$ on the set of range properties (see Equation 5) is given by Equation 12. The similarity between the two *Concepts* $A$ and $B$ on the intensional aspect is calculated as the weighted mean of $sim_d$ and $sim_r$
– **All-aspects similarity between two concepts**. The whole similarity $asim$ between two *Concepts* $A$ and $B$, from two ontology versions is computed by considering all the similarities assessed on the respective aspects involved, affected by the size of the aspect sets:

$$asim\left(A, B\right) = \frac{\sum_{a \in \mathit{\Gamma}} sim_a\left(A_a, B_a\right) \cdot \left(|A_a| + |B_a|\right)}{\sum_{a \in \mathit{\Gamma}} \left(|A_a| + |B_a|\right)} \tag{13}$$

where $A_a$ and $B_a$ are the name sets of the concepts $A$ and $B$, respectively, on the aspect $a \in \mathit{\Gamma}$, where $\mathit{\Gamma}$ is the set of all the aspect names, as defined in Equations (1)-(9), i.e., $\mathit{\Gamma}$ = $\{uri, labl, sub, sup, eq, ext, d, r\}$.
If the $asim$ value is 1, the two concepts are equal, otherwise a value in the range $[0, 1]$ describes the similarity between the concepts. The measure $asim$ can be used to analyze the drift on a concept as it changes over time, through a concept chain assembled across succeeding ontology versions. More formally, given $O^{t_1}, O^{t_2}, ..., O^{t_n}$, the $n$ successive versions of the ontology $O$, the similarity between two *Concepts* $A^{t_i}$ and $B^{t_{i+1}}$, selected from the two successive ontology versions $O^{t_i}$ and $O^{t_{i+1}}$, is assessed according to Equation 13.

### 2.3 Semantic drift assessment at ontology version level

To determine how the ontology evolves and how the semantics changes among ontology versions, the semantic drift is evaluated at the level of entire ontology versions. Comparing two ontology versions $O^{t_i} = \langle C^{ti}, R^{ti}, I^{ti}, T^{ti}, V^{ti} \rangle$ and $O^{t_j} = \langle C^{tj}, R^{tj}, I^{tj}, T^{tj}, V^{tj} \rangle$ means to find correspondences among the ontology concepts: for a concept $A^{t_i} \in C^{ti}$ in the ontology $O^{t_i}$, there must be a concept $B^{t_j} \in C^{tj}$ in $O^{t_j}$, such that the two concepts can be considered equivalent. In the Semantic Web domain, a resource is unequivocally identified by a URI (Uniform Resource Identifier); i.e., each resource has its own URI, different from any other resource. Starting from this assumption, two concepts $A^{t_i}$ and $B^{t_j}$, belonging to two different ontology versions, are considered as equal if they have the same URI (Equation 11). These concepts, with unchanged URIs across the versions, are considered in common among the versions and represented by the intersection set $|C^{t_i} \cap C^{t_j}|$. All the concepts present in the ontologies are represented as the union set $|C^{t_i} \cup C^{t_j}|$. Therefore, the semantic drift between the two ontology versions $O^{t_i}$ and $O^{t_j}$ is calculated through the overall similarity ($osim$) over the concepts from the two ontologies with the same URI. The $osim$ measure is defined as follows:

$$osim\left(O^{t_i}, O^{t_j}\right) = \frac{\sum_{\forall A^{t_i} \in C^{t_i}, \forall B^{t_j} \in C^{t_j}, A^{ti}_{uri} = B^{tj}_{uri}} asim\left(A^{t_i}, B^{t_j}\right)}{|C^{t_i} \cap C^{t_j}|} \cdot K \qquad (14)$$

where $A^{ti}_{uri}$ and $B^{tj}_{uri}$ are the URI aspects of the *Concept* $A^{t_i}$ and $B^{t_j}$, respectively; $asim$ is the all-aspects similarity between two concepts (Equation 13); $K = \frac{|C^{t_i} \cap C^{t_j}|}{|C^{t_i} \cup C^{t_j}|}$ is a value representing the ratio between the number of concepts in common among the ontologies over the number of all the individual concepts in the two ontologies. Let us notice that $K$ provides an important contribution to the similarity calculation, because it allows considering not just the concepts in common among the two ontology versions ($|C^{t_i} \cap C^{t_j}|$), but also the remaining ones ($|C^{t_i} \cup C^{t_j}|$), i.e., concepts added or removed during the ontology evolution. This way, the higher the number of concepts added or removed among the versions, the higher the semantic drift between the ontology versions.

## 3 A Case Study

This section shows the benefits of the OntoDrift methods and measures through a case study. Five consecutive DBpedia versions have been selected: *DBpedia_3_7*, *DBpedia_3_8*, *DBpedia_3_9*, *DBpedia_2015_04*, *DBpedia_2015_10*[1]. The semantic drift of the concept *Sport* among the DBpedia versions is shown in Figure 2 as a chain connecting the concepts *Sport* of different ontology versions through labels describing the similarity values calculated on concept pairs. The chain detects which version pairs have the highest drift (e.g., *DBpedia_3_8* and *DBpedia_3_9*, with $asim = 0.61$) or the lowest one (e.g., *DBpedia_2015_04*, *DBpedia_2015_10*, with $asim = 0.96$). This concept-per-concept view allows the analysis of how the concept evolves through consecutive

[1] the ontology versions are available at https://wiki.dbpedia.org/develop/datasets

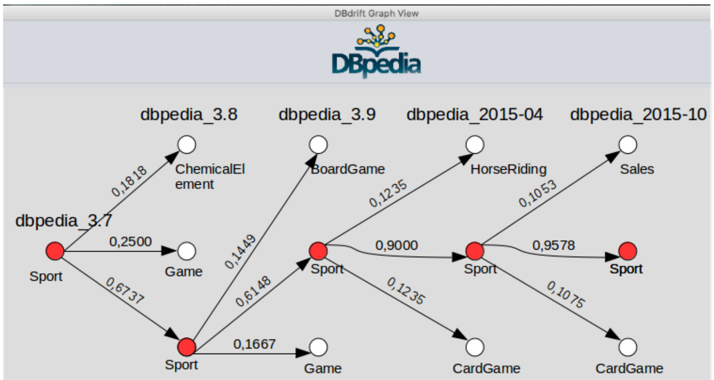


Fig. 2: The similarity *asim* on the concept *Sport* (red marked) among consecutive DBpedia versions. The other concepts are the most similar to *Sport*

versions of the ontology and provides semantic drift values. The other concepts, shown in figure, are the most similar ones to *Sport* after *Sport* itself. The semantic drift on pairs

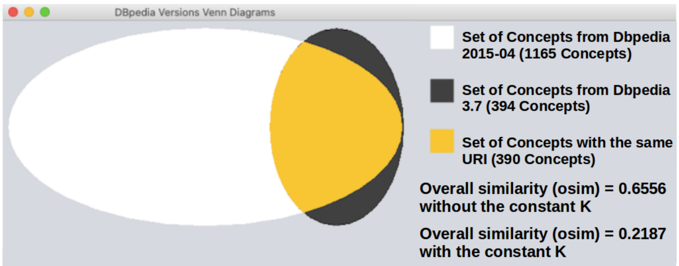


Fig. 3: The semantic drift between two DBpedia versions.

of DBpedia versions is assessed by applying the overall similarity measure ($osim$, see Equation 14). Figure 3 presents a comparison between the two versions *DBpedia_3_7* and *DBpedia_2015_04*. The Venn diagram depicts three sets: the set of concepts in *DBpedia_3_7*, the set of concepts in *DBpedia_2015_04* and the intersection set (i.e., concepts in both the versions). The identity-based solutions for the semantic drift evaluate the similarity only on the concepts in the intersection [8]. Our similarity measure $osim$, instead, includes the constant $K$ (see Equation 14) to measure the semantic drift among the versions, also considering the concepts that are not in both the versions. In fact, the drift between versions *DBpedia_3_7* and *DBpedia_2015_04* is around 34% ($osim = 0.66$) without the $K$, and around 78% ($osim = 0.22$) with the $K$. Thanks to $K$, OntoDrift-assessed drift is more accurate since it considers concepts added and removed across versions (i.e., in our case study, *DBpedia_2015_04* contains many more new concepts than *DBpedia_3_7*).

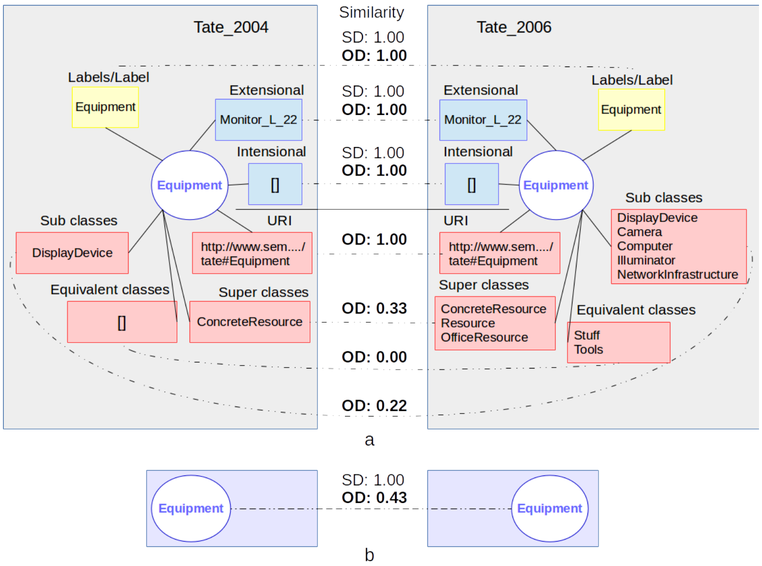


Fig. 4: OntoDrift (OD) vs. Semadrift (SD): similarity evaluation between the *Concept* Equipment of the ontology versions *Tate_2004* and *Tate_2006*.

## 4 Approach Evaluation

This section presents a comparison between OntoDrift and the framework presented in [8], called Semadrift. A two-steps comparison is given: the first one focuses on demonstrating how much OntoDrift improves the drift assessment on the single concept, whereas the second one aims at showing the effectiveness of our drift measures on entire ontology versions. The selected ontologies are Tate [8] and OWL-S profile[2], which respectively describe the cataloging of artworks and the services offered by service providers. The drift evaluated on a single concept is shown on the *Concept Equipment*, on Tate versions: *Tate_2004* and *Tate_2006*, as shown in Figure 4a. The similarity is calculated on each concept aspect from the two ontology versions by using OntoDrift (OD) and Semadrift (SD). The two approaches are compared on each concept aspect in common (in yellow) and extended (in cyan). Similarity is provided also on the new-introduced aspects (in red). The labels aspect does not change, the approaches have the same similarity on this aspect (1.0). No drift is found on the intensional and extensional aspects, that are defined in the same way in both the approaches. Similarities evaluated on new-introduced aspects, such as superclasses ($sim_{sup} = 0.33$), subclasses ($sim_{sub} = 0.22$) and equivalent classes ($sim_{eq} = 0$) highlight some changes in the ontology. According to these aspects, OntoDrift reveals a semantic drift on *Equipment* across the two versions ($asim = 0.43$, Equation 13), whereas Semadrift considers that concept unchanged ($whole\ similarity = 1$, cf.[8]), as displayed in Figure 4b. OntoDrift similarity measures can better detect any extensions or upgrades in the knowledge

[2] https://www.w3.org/Submission/OWL-S/

modeling by considering concept-related identifier and the taxonomic relations (e.g., subclass, superclass, equivalent class).

The assessment of semantic drift at ontology level is shown in Table 1 where the similarity is calculated among two consecutive ontology versions by OntoDrift considering the $osim$ measure (Equation 14), and by Semadrift through the $whole\ similarity$ measure. Let us notice that OntoDrift similarity measure causes a more sensible evaluation of the semantic drift on the entire versions. In fact, $osim$ considers more aspects than Semadrift $whole\ similarity$, including labels and taxonomic relations. OntoDrift shows weaker similarity values than Semadrift among consecutive versions of OWL-S Profile, due to the several concept taxonomic relations (i.e., some concepts are extended with subclasses, superclasses and equivalent classes) that OntoDrift evaluates. In Tate ontology, many concepts are added over time, some changes are applied on single concepts and little changes occur to relations. Since OntoDrift is quite sensitive to the concept change and extension, it returns more polished assessments on all versions. For instance, among versions going from $Tate\_2004$ to $Tate\_2013$, Semadrift assesses a stable drift (i.e., similarity in the range $[0.22, 0.25]$) while OntoDrift assesses more variable drifts (i.e., similarity in the range [0.49, 1.00]). Additionally, OntoDrift improves the identity-based approach, that considers only matching concepts across ontology versions, by evaluating the drift also on the unmatching concepts across ontology versions (see Equation 14).

Table 1: Semantic drift evaluation at ontology level

| **Compared ontology versions** | **OntoDrift** | **Semadrift** |
|---|---|---|
| OWL-S Profile 1.0 - OWL-S Profile 1.1 | 0.26 | 0.65 |
| OWL-S Profile 1.0 - OWL-S Profile 1.2 | 0.26 | 0.65 |
| OWL-S Profile 1.1 - OWL-S Profile 1.2 | 0.49 | 0.66 |
| Tate 2003 - Tate 2004 | 0.99 | 0.29 |
| Tate 2003 - Tate 2006 | 0.64 | 0.27 |
| Tate 2003 - Tate 2007 | 0.56 | 0.24 |
| Tate 2003 - Tate 2011 | 0.49 | 0.23 |
| Tate 2003 - Tate 2012 | 0.49 | 0.23 |
| Tate 2003 - Tate 2013 | 0.49 | 0.23 |
| Tate 2004 - Tate 2006 | 0.64 | 0.26 |
| Tate 2004 - Tate 2007 | 0.56 | 0.23 |
| Tate 2004 - Tate 2011 | 0.49 | 0.22 |
| Tate 2004 - Tate 2012 | 0.49 | 0.23 |
| Tate 2004 - Tate 2013 | 0.49 | 0.23 |
| Tate 2006 - Tate 2007 | 0.59 | 0.23 |
| Tate 2006 - Tate 2011 | 0.53 | 0.23 |
| Tate 2006 - Tate 2012 | 0.53 | 0.23 |
| Tate 2006 - Tate 2013 | 0.53 | 0.23 |
| Tate 2007 - Tate 2011 | 0.88 | 0.24 |
| Tate 2007 - Tate 2012 | 0.88 | 0.24 |
| Tate 2007 - Tate 2013 | 0.88 | 0.24 |
| Tate 2011 - Tate 2012 | 1.00 | 0.24 |
| Tate 2011 - Tate 2013 | 1.00 | 0.24 |
| Tate 2012 - Tate 2013 | 1.00 | 0.25 |

## 5 Conclusion

The paper presented OntoDrift, an approach to assess the semantic drift on *Concepts* among different ontology versions. The approach provides a novel definition of *Concept*, which includes a wide set of related features, called aspects. Similarity measures are defined to assess the semantic drift among concepts and ontology versions by considering the multiple-aspect concept definition. The benefits of the approach are various, first of all, the semantic drift assessment is more accurate, because it is evaluated on multiple aspects, not only including concept labels, intension and extension, but also the URIs and taxonomic relations. The method can be used to assess the drift among ontology versions and knowledge graphs (e.g., DBpedia), thanks to the identity-based approach design. Additionally, the indentity-based approach is extended to consider not only the concepts in common among ontology versions, but also those added and removed during the ontology evolution to provide more refined drift assessments.

## References


1. Burek, P., Scherf, N., Herre, H.: A pattern-based approach to a cell tracking ontology. Procedia Computer Science **159**, 784 – 793 (2019), Knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 23rd International Conference KES2019
2. Cavaliere, D., Loia, V., Senatore, S.: Towards an ontology design pattern for uav video content analysis. IEEE Access **7**, 105342–105353 (2019)
3. Fanizzi, N., d'Amato, C., Esposito, F.: Conceptual clustering and its application to concept drift and novelty detection. In: Proceedings of the 5th European Semantic Web Conference on The Semantic Web: Research and Applications. pp. 318–332. ESWC'08, Springer-Verlag, Berlin, Heidelberg (2008)
4. Frermann, L., Lapata, M.: A bayesian model of diachronic meaning change. Transactions of the Association for Computational Linguistics **4**, 31–45 (12 2016)
5. Gulla, J., Solskinnsbakk, G., Myrseth, P., Haderlein, V., Cerrato, O.: Semantic drift in ontologies. In: WEBIST 2010. vol. 2, pp. 13–20 (01 2010)
6. Hamers, L., Hemeryck, Y., Herweyers, G., Janssen, M., Keters, H., Rousseau, R., Vanhoutte, A.: Similarity measures in scientometric research: The jaccard index versus salton's cosine formula. Information Processing & Management **25**(3), 315 – 318 (1989)
7. Hamilton, W.L., Leskovec, J., Jurafsky, D.: Diachronic word embeddings reveal statistical laws of semantic change. In: Proc. of the 54th Annual Meeting of the Association for Computational Linguistics (Vol. 1: Long Papers). pp. 1489–1501 (2016)
8. Stavropoulos, T., Andreadis, S., Kontopoulos, E., Kompatsiaris, I.: Semadrift: A hybrid method and visual tools to measure semantic drift in ontologies. Journal of Web Semantics **54**, 87 – 106 (2019)
9. Wang, S., Schlobach, S., Klein, M.: Concept drift and how to identify it. Journal of Web Semantics **9**(3), 247 – 265 (2011)
10. Wittek, P., Daranyi, S., Kontopoulos, E., Moysiadis, T., Kompatsiaris, I.: Monitoring term drift based on semantic consistency in an evolving vector field (07 2015)