

Temporalité et graphes de connaissances : analyse théorique et enjeux pratiques

W. Charles¹, N. Aussenac-Gilles¹, N. Hernandez^{1,2}

¹ IRIT- CNRS et Université de Toulouse, prenom.nom@irit.fr

² Université Toulouse 2 Jean Jaurès

19 mai 2023

Résumé

La représentation de faits ancrés dans une temporalité est une tâche ardue en RDF. Nous détaillons les étapes menant à la conception d'une ontologie intégrant une temporalisation des faits. Nous dégagons notamment un cadre théorique permettant de déterminer quelles propriétés temporaliser selon la conception du temps retenue, ainsi que des raisons pouvant pousser à simplifier celui-ci. Nous revenons également sur les différentes méthodes de représentation du temps et de la temporalisation des faits.

Mots-clés

Temporalité, Ontologie, Analyse Théorique, Philosophie.

Abstract

Representing temporally anchored facts in RDF is a challenging task. Here, we detail steps that can help designing an ontology integrating temporal facts. We notably propose a theoretical reference frame aiming to identify properties that are to be temporalised, as well as motives that could lead to simplifying it. We also review the various ways to represent time and fact temporalisation.

Keywords

Temporality, Ontology, Theoretical Analysis, Philosophy.

1 Introduction

Dans le contexte du Web Sémantique, RDF s'est imposé au cours des dernières années comme le standard de la représentation de connaissances [24]. Ce langage reposant sur l'usage de triplets permet de mettre en relation deux ressources qualifiées par un identifiant unique (URI) au travers d'un prédicat possédant lui-même un URI. Bien qu'extrêmement simple, ce formalisme offre de vastes possibilités de représentation. Des vocabulaires tels que OWL ou RDFS proposent une extension de la sémantique de celui-ci et accroissent son expressivité et ces possibilités de raisonnement [38]. Toutefois, si RDF permet de représenter sans difficulté des relations entre entités à un instant donné, la représentation d'entités à composante temporelle, ou plus généralement de faits temporellement dépendants y est autrement plus difficile [5]. En effet, le formalisme du triplet ne permet par exemple pas d'apposer un marqueur tempo-

rel sur une relation sans l'usage d'un mécanisme auxiliaire. De nombreuses solutions ont été mises en avant pour pallier ce manque, allant de l'usage de formalismes utilisant la sémantique RDF tels que la réification [21], jusqu'à des extensions de la syntaxe en utilisant par exemple des quadruplets au lieu de triplets [45].

Si cette limite pratique peut encore constituer un obstacle à l'intégration de données temporelles dans un graphe de connaissances, d'autres questionnements sous-tendent la représentation temporelle. En effet, dès lors que l'on commence à ajouter des composantes temporelles, il devient nécessaire de s'interroger sur la légitimité de temporaliser telle ou telle entité ou propriété. Pour ce qui est des entités, des travaux tels que ceux de l'ontologie de haut niveau DOLCE [16] introduisent deux grandes catégories d'entités temporelles : les endurants et les perdurants (similairement aux occurrents et continuants de BFO, une autre ontologie de haut niveau [4]), distinguables par la stabilité au cours de leur existence des uns, et la propension à connaître plusieurs états au cours de celle-ci des autres. Notons que cette distinction exclut *a priori* l'existence d'entités hors du temps. Or, les ontologies pratiques faisant usage de la temporalité mêlent souvent des entités temporellement marquées à d'autres qui ne le sont pas. On peut citer comme exemple l'ontologie des organisations¹, une recommandation du W3C. Bien que celle-ci inclue une dimension temporelle dans la description des organisations et des rôles au sein de celles-ci, elle n'en intègre pas pour la description des sites occupés par les entreprises. Cette pratique vise à simplifier la représentation dans son ensemble et permet de limiter le coût en stockage. Elle se justifie par l'intérêt limité que l'on peut porter à certaines ressources limitrophes au contexte visé par l'ontologie, ou par l'immutabilité locale de certains objets dans le cadre d'une représentation ciblée. La description d'entités temporelles n'est cependant pas complète sans la temporalisation des propriétés qui les lie. Toutefois, s'il semble naturel de temporaliser certaines relations tel que l'union conjugale (*estMariéeA*), d'autres telles que la relation de paternité (*perDe*) semblent plus délicates. De par son usage courant, il serait tentant de considérer *perDe* comme une relation atemporelle. On pourrait toutefois objecter que cette propriété n'entre en

1. <https://www.w3.org/TR/vocab-org/>

existence qu'à l'instant où l'enfant cible de la relation vient au monde.

Dans cet article, nous tâcherons d'analyser les étapes et les choix de représentation à suivre lors de la conception d'une ontologie pratique intégrant une dimension temporelle. Il ne s'agit pas de donner une méthodologie définitive, mais simplement de souligner les complexités sous-jacentes à chaque étape. L'accent sera mis sur les aspects de représentation et non sur ceux de raisonnement. La section 2 propose des distinctions théoriques à réaliser lors des choix de temporalisation d'entités et de propriétés. Les considérations issues de cette section constituent la contribution principale de cet article. La section 3 traitera les aspects inhérents à la représentation du temps lui-même. La section 4 s'attachera à dresser un état de l'art des techniques employées afin de représenter la temporalité à l'aide de RDF.

2 Temporalité : que représenter ?

Avant de s'interroger sur comment représenter le temps, il est nécessaire de définir quelles ressources manipulées au sein d'une représentation doivent être ancrées dans le temps. Concernant les entités, la distinction entre durables et perdurants, issue de considérations philosophiques classiques [32], est portée notamment par l'ontologie de haut niveau DOLCE [16]. Les entités perdurantes sont décomposées en un ensemble de parties temporelles correspondant à ce qu'est un individu à un instant donné. A un instant donné, seule une partie temporelle de l'entité perdurante est présente, ce qui signifie qu'elle n'est pas pleinement présente. Les entités durables, au contraire, sont pleinement présentes à tout instant de leur existence. Cela ne signifie pas qu'elles sont hors du temps, car elles peuvent avoir une période d'existence fixée. En termes concrets, il s'agit d'une distinction entre entités qui évoluent (perdurants) et qui n'évoluent pas (durables).

Il reste toutefois à s'interroger sur la temporalisation des triplets mettant en relation ces entités. Nous inscrivons ici dans un cadre théorique les choix de représentation induisant la temporalisation de triplets. Ce cadre se fondant sur des considérations d'ordre philosophique, il est vraisemblable que la représentation pratique du temps dans le contexte d'une ontologie de domaine induise des simplifications. Dans un second temps, donc, nous mettons en exergue les raisons qui peuvent amener à simplifier ce cadre dans la pratique.

2.1 Considérations théoriques

L'interrogation qui nous occupe est la temporalité de propriétés liant deux entités. En nous appuyant sur l'exemple de pereDe , nous proposons une analyse des choix de représentation qui peuvent mener ou non à temporaliser la propriété. Deux dimensions sont ici considérées. La première est l'ontologie du temps retenue. On ne parle pas d'ontologie pratique mais bien d'ontologie au sens de qualification de la nature du temps. Ainsi, il existe principalement dans la philosophie trois grandes ontologies du temps [37, 8]. *L'éternalisme* affirme que le passé, le présent et le futur ont une réalité éternelle. Le *présentisme* considère

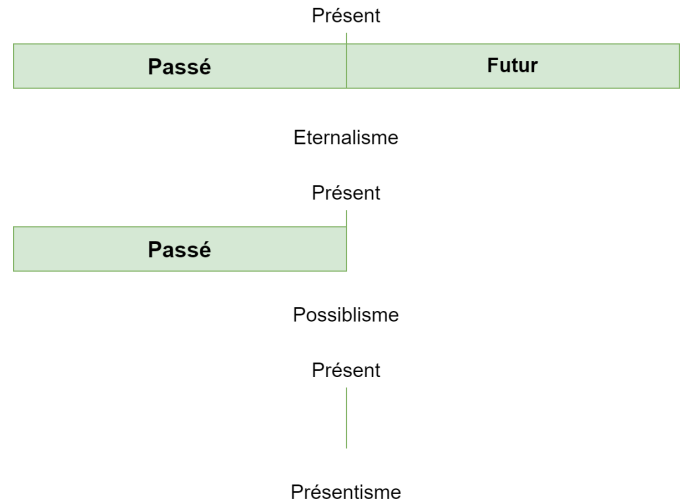


FIGURE 1 – Trois visions du temps

tout ce qui n'est pas présent est fiction, et s'appuie sur l'intuition cognitive naturelle qu'à l'individu de l'instant présent. Enfin, le *possibilisme* [37] (aussi appelé *gradualisme* [43] ou *théorie du bloc croissant* [44]) considère que le futur n'existe pas encore. Le passé y est conceptualisé comme l'ensemble des présents déjà parcourus. La figure 1 résume visuellement les composantes temporelles considérées comme existantes dans chacune des trois théories.

La seconde dimension est le monde conceptuel dans lequel s'inscrit la relation. Plus explicitement, la relation concerne-t-elle l'entité en tant que concept ou en tant qu'objet physique ? Nous allons donc analyser le statut de la propriété pereDe dans chacune des conceptions temporelles, selon que l'on considère qu'elle lie deux entités conceptuelles ou physiques. Dans ce paragraphe, on parlera de propriété *atemporelle* pour une propriété qui n'a pas vocation à être contextualisée dans le temps, de propriété *temporelle* pour une propriété qui n'est valable que sur des périodes temporelles de durée finie. Enfin, on parlera de propriété *semi-temporelle* pour celles dont la validité a un début, mais pas de fin.

2.1.1 Présentisme

Dans le contexte du présentisme, le passé et le futur n'ont pas de réalité. Ainsi, la notion de temporalisation n'a de sens qu'en tant que fiction mathématique [8]. Dans un contexte strictement présentiste :

- Si pereDe lie deux entités conceptuelles : la propriété pereDe est atemporelle.
- Si pereDe lie deux entités physiques : la relation ne peut exister que si les entités existent. Autrement dit, dans le cas de d'un humain, la relation n'existera que si les deux individus qu'elle lie sont présentement vivants. La propriété est donc atemporelle.

On notera que dans une vision présentiste, on peut néanmoins discuter sur des entités passées en considérant qu'elles correspondent à un univers fictif suivant une représentation possibiliste. Les entités représentées ne sont alors plus les entités réelles mais des projections dans un univers

Modèle temporel	Entités Conceptuelles	Entités Physiques
Présentisme	Atemporelle	Atemporelle si existante
Possibilisme	Semi-temporelle	Temporelle ou Semi-Temporelle
Eternalisme	Atemporelle	Temporelle

TABLE 1 – Nature des propriétés selon l’ontologie du temps et l’aspect des entités considérés

fictif.

2.1.2 Possibilisme

Le possibilisme établit l’existence du passé et du présent. Dans ce contexte :

- Si pereDe lie deux entités conceptuelles : la propriété pereDe est semi-temporelle. En effet, il existe un présent précédant la naissance de l’individu où même le concept de l’individu n’existe pas. La propriété pereDe n’est donc valable qu’à compter du moment où la conceptualisation de la réalisation de la propriété est possible.
- Si pereDe lie deux entités physiques : la propriété est temporelle, ou semi-temporelle selon que les individus concernés soient tous deux en vie ou non.

2.1.3 Eternalisme

L’éternalisme considère l’existence perpétuelle du passé, présent et futur. Nous en tirons les conséquences suivantes :

- Si pereDe lie deux entités conceptuelles : la propriété est atemporelle. En effet, le futur existant, la relation conceptuelle existe également à tout instant.
- Si pereDe lie deux entités physiques : la propriété est temporelle, car elle lie deux entités physiques qui n’existent qu’à une période donnée. On ne peut pas toujours savoir à quel instant elle prendra fin, mais elle demeure temporelle car le futur est déjà partie intégrante de la réalité.

La table 1 récapitule la section 2.1.

2.2 Considérations pratiques

En pratique toutefois, ces considérations sont parfois écartées de sorte à obtenir une représentation davantage adaptée aux besoins du domaine représenté. Nous avons identifié deux cas distincts permettant de justifier l’omission de la temporalisation de certaines entités, tous liés à la notion de domaine de représentation.

2.2.1 Notion limitrophe au cadre de l’ontologie

Le premier cas englobe les situations où la temporalité est supprimée de sorte à ne pas alourdir la représentation plus que nécessaire. Formellement, pour un triplet (s, p, o) , légitimement temporel dans le cadre de représentation retenu par l’ontologie, on distingue plusieurs options ;

- Le triplet n’est pas temporalisé, et o non plus. C’est le cas de l’ontologie des organisations, mentionnée dans l’introduction. Ce cas appauvrit la représentation temporelle de s , mais cela peut être motivé

par le peu d’importance que l’on accorde à certaines propriétés.

- Le triplet est temporalisé, mais o ne possède aucune représentation temporelle. Dans le contexte de la représentation de territoires évolutifs par exemple, il est possible de représenter un lien temporellement marqué avec l’individu régissant un territoire, sans pour autant donner une représentation temporelle à l’individu en question. La représentation de celui-ci n’est pas ce qui nous occupe dans ce contexte. Bien que ne limitant pas la représentation temporelle de s , cette solution complique la réutilisabilité dans le cas de la confrontation à un jeu de données inscrivant o dans une temporalité (typiquement si le jeu de données fait usage d’un modèle *fluents* décrit dans la section 4.2.1).
- Le triplet n’est pas temporalisé, mais o l’est. Ce cas de figure peut se présenter quand la propriété p n’a guère d’importance dans notre contexte. Dans des travaux précédents [12], nous avons également mis en avant l’usage de cette solution pour alléger la représentation en temporalisant implicitement la propriété p . Cette dernière était alors considérée valide à l’intersection des intervalles marquant l’existence des objets temporels o et s .

2.2.2 Immuabilité locale au cadre d’étude

Certaines entités peuvent être considérées comme localement immuables du fait du contexte retenu pour l’ontologie. Par exemple, dans le contexte d’une ontologie de représentation de territoires contemporains [15], il est possible de considérer les éléments topographiques tels que les montagnes comme des éléments immuables. En effet, bien qu’ils évoluent au fil du temps et soient voués à disparaître, la période couverte par le discours ne justifie par leur temporalisation au sein de l’ontologie.

3 Représentation du temps

Les considérations décrites dans la section précédente permettent de trancher quant aux composantes de l’ontologie susceptibles d’être plongées dans le temps. Avant de représenter la temporalité de celles-ci, il convient toutefois de s’interroger sur le formalisme temporel à retenir pour représenter le temps en lui-même.

3.1 Représentation théorique du temps

D’un point de vue ontologique, plusieurs approches ont été mises en avant pour représenter et raisonner sur le temps. Cette représentation est polarisée entre les représentations à base de points et celles à base d’intervalles, les premières étant fondées sur l’intuition mathématique et physique du temps, tandis que l’autre repose davantage sur l’expression du temps en langage naturel [3].

3.1.1 Le temps en points

La première option consiste à considérer le temps comme un ensemble ordonné de points [9, 31]. Cette approche permet de représenter des intervalles temporels, soit à l’aide d’ensembles de points [31] (on discrétise alors le temps),

soit à l'aide de paires (a, b) de points ordonnés formant un intervalle $[a, b]$. Dans ce dernier cas, l'intervalle est matérialisé par un instant de début a et un instant de fin b [9]. Bien que ces approches aient leurs intérêts qui justifient leur usage aujourd'hui encore (cf Table 2), elles peinent à représenter les *Points de rupture* [29], instants auxquels une propriété change de valeur logique. En effet, supposons une propriété t vraie sur un intervalle $[p1, p]$ et fausse sur un intervalle $[p, p2]$. Que se passe-t-il alors au point p ? La propriété est-elle vraie? Fausse? Vraie et fausse? Ni vraie ni fausse?

3.1.2 Intervalles temporels

Cette limite des modèles à base de points pousse à la mise en place d'un modèle basé sur des intervalles. L'algèbre temporelle d'Allen [1] définit un ensemble de 13 relations permettant de décrire et raisonner sur des intervalles. Cette représentation est fondée davantage sur la représentation des liens entre intervalles que sur un ancrage temporel de ceux-ci. Autrement dit, le raisonnement à l'aide d'intervalles d'Allen s'effectue à l'aide des relations définies par Allen et des lois de composition proposées, et ne nécessite pas d'imposer des dates à ceux-ci. Ainsi, pour le problème précédent, on a alors t vraie sur un intervalle $i1$ et fausse sur un intervalle $i2$ avec $meets(i1, i2)$ (signifiant que les intervalles $i1$ et $i2$ sont directement consécutifs). Le problème du point de jonction est évacué. Toutefois quand on parle d'intervalle ici, il faut bien comprendre qu'on ne parle pas de l'objet mathématique dont on a l'habitude, mais d'un objet défini par la sémantique de l'algèbre d'Allen. Ainsi, la version initiale ne rend pas possible la représentation d'intervalles de durée nulle (ex : $[a, a]$). Cet aspect est argumenté par le fait qu'en pratique, on peut ramener un instant à un intervalle de temps de durée très faible. Cet argument est battu en brèche dans le contexte de l'application des lois de la physique [29]. Le contre-exemple classique est la description de l'instant où la vitesse d'une balle lancée en l'air est nulle, que la physique identifie comme étant un instant et non un intervalle de durée très faible. Pour pallier ce manque, Allen proposera plus tard [2] la notion de moment qui constitue un intervalle insécable, ainsi que la notion de nid, qui permet de définir des concepts équivalents aux points en mettant en avant une relation d'ordre complète sur des ensembles d'intervalles. Toutefois, cette représentation est ardue à utiliser en pratique.

3.1.3 Une combinaison points et intervalles

[29] tente d'unifier le meilleur des deux mondes en proposant une théorie du temps mélangeant points et intervalles. Il met en place un système mathématique à base de relations étendant celles de l'algèbre d'Allen. A l'instar de celle-ci, le modèle proposé consiste davantage en une représentation fondée sur le langage naturel et les relations liant les objets temporels qu'en une représentation du temps traditionnelle faisant usage de marqueurs temporels indexables par des entiers.

3.2 Représentation pratique du temps

Dans le cadre des recherches sur l'intelligence artificielle, de nombreux modèles de représentation temporelle ont été proposés. Les formalismes présentés dans la section précédente ne s'attachent toutefois qu'à décrire les relations entre objets temporels.

3.2.1 Intervalles mathématiques

En pratique, de nombreux modèles font usage d'intervalles mathématiques classiques. Dans le contexte des modèles à base de points, un intervalle I est défini par deux points a et b ($a < b$). Certains modèles [6, 25, 19, 14] utilisent simplement ce formalisme. Or, en mathématiques, on distingue pour les intervalles des bornes fermées et ouvertes, indiquant si la borne concernée appartient ou non à l'intervalle. Notons que cette approche présente un défaut dans un contexte de représentation informatique. En effet, la représentation machine implique nécessairement une discrétisation du temps, tandis que la qualification des bornes évoquée précédemment se fonde sur la représentation d'intervalles continus. Néanmoins, cette représentation est utilisée en pratique, (parfois en conjonction avec l'algèbre d'Allen [25]), car elle permet notamment de lever le problème du point de rupture. Dans la pratique on distingue deux cas d'usage :

- La spécification des bornes est libre [6, 25, 19]. Il est donc de la responsabilité de l'utilisateur de construire un modèle cohérent.
- Une borne est fixée comme étant ouverte (par exemple la borne supérieure dans [14]). Notons que dans ce cas, le cadre de représentation tranche le problème du point de rupture, en indiquant qu'en ce point la propriété prend la valeur qu'elle aura par la suite.

3.2.2 Ontologies pratiques de représentation du temps

En pratique, la représentation du temps dans les ontologies est souvent réalisée à l'aide de OWL-Time². Sur 19 ontologies du temps (portant le tag "Time") présentes sur Linked Open Vocabularies (LOV³) au moment de la rédaction de cet article, OWL-Time est de loin la plus utilisée avec 50 liens entrants (cf figure 2). Cette ontologie constitue un standard de représentation du Web Sémantique. Elle couvre une vaste palette de possibilités de représentations explicites du temps, aussi bien du point de vue d'une représentation proche d'usages usuels (notion de jour, mois, années, calendrier ...) que mathématique (usage de points temporels pour ancrer les éléments dans le temps). Elle propose également une implémentation de l'algèbre d'Allen pour les intervalles. OWL-Time fait notamment usage des littéraux `xsd:date` et `xsd:dateTime` qui sont parfois utilisés directement au sein d'ontologies afin de représenter certains aspects temporels sans le truchement d'OWL-Time [41]. Toutefois le défaut principal d'OWL-Time est que les primitives proposées ne permettent pas de représenter explicitement le flou, i.e. l'imprécision et/ou l'incertitude liées à

2. <https://www.w3.org/TR/owl-time/>

3. <https://lov.linkeddata.es/dataset/lov>

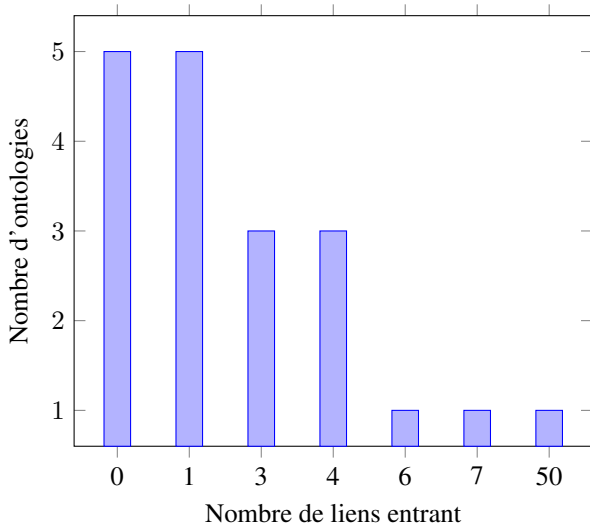


FIGURE 2 – Répartition du nombre de liens entrants pour les ontologies portant le tag "Time" référencées sur le LOV

cette dimension.

3.2.3 Incertitude et imprécisions temporelles

Dans OWL-Time la seule représentation d'une quelconque incertitude ou imprécision est implicite. En effet, lors de la spécification d'un `time:Instant`, la valeur précisée peut être simplement un mois ou une année, permettant de maintenir une imprécision quant à la date effective de l'instant en question. Nombre de modèles plus expressifs existent. Tout d'abord, il existe des modèles reposant sur la pondération de valeurs temporelles, associés à un modèle de raisonnement sur des données incertaines (exemple : logique de Markov [25]). Intuitivement, cette pondération peut être assimilée à la probabilité qu'un fait se soit produit à une date donnée. Toutefois, que ce soit pour [25], [45] ou encore [14], la représentation d'une temporalité vague ne porte pas sur la composante temporelle, mais sur le triplet temporalisé (cf section 4.1). Il est par exemple impossible de déclarer qu'on est certain qu'un fait ait eu lieu sans avoir une idée de sa date exacte.

Comme mentionné précédemment, la représentation de l'algèbre d'Allen repose davantage sur les relations entre intervalles que sur leur datation. Ainsi, ces relations peuvent être utilisées pour représenter certains types d'informations temporelles vagues. Par exemple, elles permettent de décrire des intervalles vagues comme "*Un moment X dans l'après-midi.*" : `during(ApresMidi, X)` sans avoir à définir précisément ce qu'est l'après-midi (c'est simplement un intervalle après le matin, et avant le soir). Elles ne permettent cependant pas de représenter des périodes telles que "*Un moment X aux alentours de 13h*" sans définir un intervalle des "alentours".

Une approche commune dans le contexte des humanités numériques, car recommandée par le modèle conceptuel CIDOC-CRM [10] est l'intervalle à quatre points. La borne de début d'intervalle d y est remplacée par une paire (d_{min}, d_{max}) formant un intervalle dans lequel

début effectif de la période décrite se trouvera. Il en va de même pour la borne de fin. Concernant "*Un moment X aux alentours de 13h*", l'intervalle à quatre points permet d'être moins précis quant à la définition de l'intervalle des "alentours", mais requiert toujours une définition de celui-ci.

Dans ce même contexte, on trouve également des ontologies ou des *gazeteers* qui utilisent pour marquer temporel des périodes déterminées par des civilisations [18] ou géologiques [13]. Une autre forme de temporalité vague consiste à représenter des nombres d'occurrences au cours d'une période donnée. Cette méthode est appelée *triplet indéterminé* dans [39]. La partie temporelle du modèle SEAS [28] met également en avant ce type de représentations, accompagnée de méthodes permettant d'effectuer des calculs sur les occurrences.

Enfin, Temporal OWL propose la notion de temporalité anonyme qui permet d'indiquer qu'un fait se produit à un instant donné sans pour autant devoir spécifier une quelconque information sur la période en question [22].

3.2.4 Autres marqueurs spéciaux

La représentation de données temporelles se heurte au fait que les connaissances que l'on peut avoir s'arrêtent souvent à l'instant présent. Ainsi, plusieurs formalismes proposent la mise en place d'un marqueur matérialisant l'instant présent [6, 21, 39, 33, 19]. On trouve globalement trois marqueurs pour réaliser ceci : `NOW`, `UC` (pour *Until Changed*) et ∞ . Si certains modèles comme Temporal RDF font emploi de plusieurs marqueurs auxquels ils attribuent des sémantiques distinctes, d'autres [19] en utilisent un seul pour couvrir plusieurs sens. Il est donc difficile de donner une sémantique universelle à ces trois marqueurs. Comme souligné par [19], toutefois, ces marqueurs sont hérités des bases de données relationnelles temporelles, où `NOW` et ∞ sont utilisés afin de représenter des connaissances valables jusqu'à l'instant présent ou à l'infini, tandis qu'`UC` est utilisé dans un contexte transactionnel. Le symbole ∞ , issu des mathématiques, est plus délicat car il est utilisé pour représenter alternativement l'instant présent ou le fait que l'intervalle de validité d'une propriété est étendu jusqu'à l'infini.

4 Représenter la temporalité

Enfin, afin de temporaliser les faits, il est nécessaire de choisir un formalisme de représentation pour les graphes de connaissances adapté aux besoins de temporalisation identifiés selon les principes énoncés section 2. Nous dressons ici un état de l'art des méthodes utilisés afin de lier les ressources aux marqueurs temporels issus d'un formalisme de la section 3. Un état de l'art plus détaillé est proposé dans [46]

4.1 Marquage temporel des propriétés

Une approche courante pour représenter la temporalité au sein d'un graphe RDF consiste à apposer un marqueur temporel à un triplet, indiquant sa période de validité [6, 21, 49]. Toutefois, le formalisme RDF ne permet pas à

lui seul cette représentation. Des mécanismes conçus pour l'ajout de métadonnées peuvent alors être utilisés [45]. La figure 3 présente la représentation temporelle de deux triplets ayant la même période de validité à l'aide de divers mécanismes que nous allons détailler ici.

4.1.1 Réification standard et Temporal RDF

La réification standard⁴ est un mécanisme introduit par le W3C permettant de représenter des méta-propriétés [30]. Un triplet y est représenté sous la forme d'une instance de la classe `rdf:statement` auquel on adjoint les composantes du triplet à l'aide des propriétés `rdf:subject`, `rdf:predicate` et `rdf:object`. La figure 3a présente un exemple de réification. L'une des premières approches dédiées à la représentation temporelle dans RDF, Temporal RDF [21, 22], fait appel au mécanisme de réification, en utilisant un vocabulaire qui lui est propre. Cette approche traite la temporalité en adjoignant à tout triplet (s, p, o) une composante temporelle t indiquant la période de validité du triplet. En pratique nombre de formalismes de temporalisation des triplets font usage de la réification RDF [6, 25] bien que certains définissent en sus une syntaxe propre à leur modèle de représentation temporelle, complétée d'un formalisme permettant la conversion vers le modèle RDF en faisant usage de la réification.

4.1.2 Réification standard : alternatives

Le défaut majeur de la réification RDF standard est sa lourdeur syntaxique, qui amène souvent les auteurs à définir leur propre syntaxe afin d'alléger la représentation. Ainsi, de nombreuses alternatives à la réification standard ont été mises en avant, chacune présentant des avantages et des inconvénients dans le cadre de la représentation temporelle.

- *Les relations n-aires*⁵ permettent de représenter des relations de cardinalité supérieure à la binarité imposée par le triplet. Elles consistent à matérialiser la relation par une entité à laquelle sont rattachées diverses propriétés. Bien que plus légère syntaxiquement que la réification, l'utilisation de relations n-aires nécessite la définition préalable de classes dont les entités matérialisant les relations seront des instances, repoussant une partie de la complexité dans la conception de l'ontologie associée. Dans le contexte de l'adjonction d'un marqueur temporel, deux options d'utilisation des relations n-aires sont possibles. La première, illustrée figure 3b, consiste à créer une classe de relation pour chaque propriété à laquelle on souhaite adjoindre un marqueur temporel. La seconde, illustrée figure 3c, ne crée qu'une seule classe de relation qui correspond aux propriétés temporelles, sur laquelle viendront se greffer l'ensemble des propriétés. Ainsi, plusieurs relations vérifiées à une même période seront regroupées sur une unique relation n-aire. Cette seconde utilisation, bien que plus compacte pour le stockage, semble peu intuitive, et présente également le défaut de détourner les relations n-aires de leur sémantique première.

En effet, là où les relations n-aires escomptent une valeur pour chacune des propriétés définies pour la classe des entités représentant la relation, le nombre de valeurs effectivement renseignées est ici variable.

- *Les graphes nommés* [11] sont un formalisme visant à permettre de faire référence à un ensemble de triplets, regroupés au sein d'un graphe, comme ressource RDF (cf figure 3d). Dans le contexte de la temporalisation, ce mécanisme n'a d'intérêt que si un grand nombre de triplets partagent la même période de validité. La représentation de tels graphes implique que dans le cas où un triplet serait valide à des périodes non connexes, celui-ci devrait être dupliqué dans chacun des graphes temporellement marqués pour lequel il est vérifié.
- *RDF-star* [23] est une extension de RDF visant à permettre de considérer un triplet comme une ressource (figure 3e). Il s'agit d'une forme de réification alternative à celle définie dans la sémantique de RDF, qui s'accompagne de syntaxes simplifiées pour alléger l'écritures de requêtes et de graphes de connaissances. Bien que naturel pour l'ajout d'une composante temporelle à un triplet, ce formalisme est encore peu utilisé. Sa standardisation par le W3C est toutefois en cours.
- *Les Singleton Properties* [34] (figure 3f) sont un mécanisme consistant à utiliser une réalisation unique de la propriété dans chaque triplet. Ainsi, si on utilise une propriété p dans deux triples différents, ces derniers utiliseront en pratique des propriétés $p\#1$ et $p\#2$. Ces propriétés singletons étant utilisées dans un unique triplet, elles peuvent ainsi servir à le désigner. Une composante temporelle peut donc par exemple y être apposée. A noter qu'une extension des premiers travaux [35] permet de préserver l'inférence lors de l'usage de propriétés singletons.
- Enfin, les méthodes d'annotation telles que aRDF [45] permettent l'ajout de la composante temporelle sous forme d'annotation. Il existe notamment des formalismes d'annotation, tels que RDFt [48], qui se concentrent exclusivement sur l'annotation temporelle. Toutefois, ces mécanismes reposent sur des extensions de la syntaxe RDF, ce qui complique leur réutilisabilité.

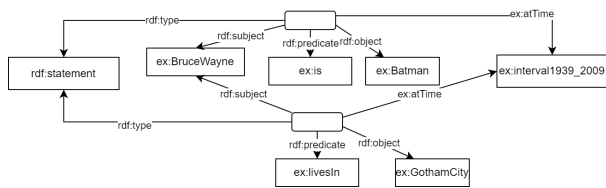
Il est à noter que nombre de formalismes de représentation du temps [6, 21, 33, 25] fondent leur sémantique sur un formalisme abstrait reposant sur l'adjonction d'une composante au triplet, certains allant même jusqu'à s'abstraire de toute implantation [33]. Ainsi, si les réalisations proposées s'appuient sur la réification, il semble pertinent de considérer des alternatives à celle-ci dans le but de simplifier le stockage [5].

4.2 Représentation de séries temporelles

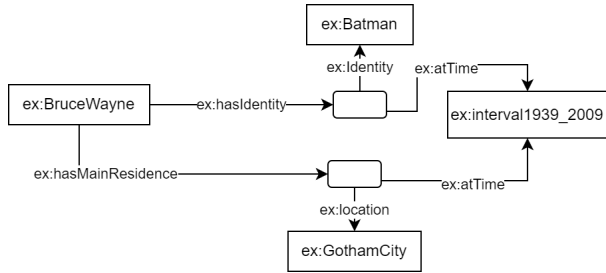
L'estampillage temporel de triplets constitue une solution pour représenter des connaissances ancrées dans une temporalité fixée. Une autre solution consiste à porter la com-

4. <https://www.w3.org/wiki/RdfReification>

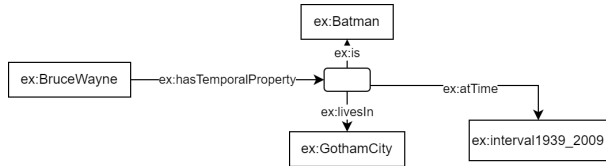
5. <https://www.w3.org/TR/swbp-n-aryRelations/>



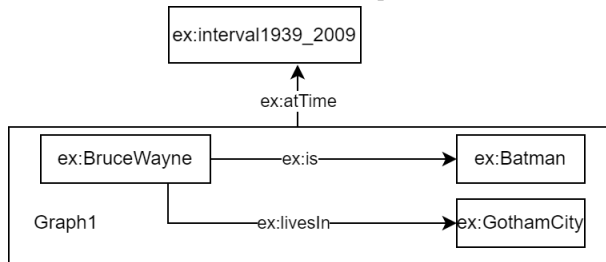
(a) Réification standard



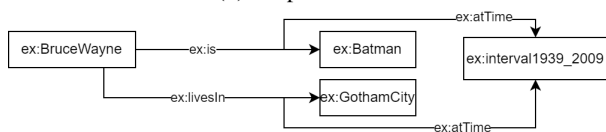
(b) Relation n-aire : option 1



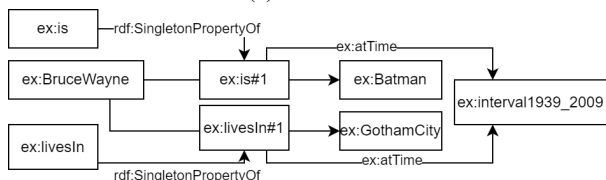
(c) Relation n-aire : option 2



(d) Graphe nommé



(e) RDF-star



(f) Singleton Property

FIGURE 3 – Réification standard et alternatives pour représenter la temporalité

posante temporelle non sur les triplets mais sur les entités représentées. On parle de sérialisation, car les entités sont alors représentées par une série d'états temporels. [20] propose deux modèles de sérialisation : les ontologies SPAN et SNAP. Les premières décrivent le cas où chaque entité déroule ses parties temporelles au cours du temps dans un unique graphe de connaissances. La seconde correspond au cas où l'on décrit une série de "snapshots" représentant l'état complet des connaissances à un instant donné.

4.2.1 Le modèle fluents

Le modèle fluents [47] est mis en avant pour représenter des entités perdurantes. Il repose sur un mécanisme relativement simple consistant à associer à chaque entité plusieurs parties temporelles, représentant son état à diverses périodes. Cette approche a par ailleurs été étendue afin de représenter une entité dans un contexte de nature quelconque (pas nécessairement temporel) [17]. Toutefois, de nombreuses questions sont soulevées en pratique. Tout d'abord, on notera que, selon DOLCE, des entités perdurantes sont amenées à coexister avec des entités endurantes, ce qui peut complexifier le graphe. Une autre complexité porte sur l'usage de propriétés que l'on qualifiera d'atemporelles. La date de naissance d'un individu, par exemple, ne dépend pas de son état, et pourra ainsi être attachée non pas à chacune des parties temporelles, mais bien à l'entité porteuse de l'identité de l'individu. Enfin, outre l'expression de requêtes qui est considérablement alourdie, le principal problème généré par cette approche est la prolifération de triplets, comme avancé par [5]. Tout d'abord, la notion de partie temporelle implique que l'ensemble des propriétés permettant de qualifier un état doivent être décrites pour chacun des états. Si la description porte sur un être humain, et que les parties temporelles donnent sa taille et son nom, pour chaque évolution de la taille, il sera nécessaire d'ajouter un nouveau triplet rattachant la nouvelle partie temporelle au nom de l'individu même si celui-ci reste identique. Un autre facteur de prolifération est la mise en relation entre diverses entités perdurantes. En effet, si une partie temporelle de A est en lien avec B , mais que sur la période couverte par ladite partie de A , B possède plusieurs parties temporelles, il peut être nécessaire de fragmenter la partie temporelle de A pour être en adéquation avec B . Cette complexité peut être réduite en considérant que la validité d'une relation entre deux parties temporelles se trouve à l'intersection des périodes de validités de celles-ci, comme mis en avant par des travaux précédents [12].

4.2.2 Séries de graphes

L'usage de séries de graphes pour représenter la temporalité consiste à représenter l'ensemble des connaissances valables à un instant donné sous forme d'un graphe de connaissances atemporel. L'ensemble des graphes ainsi produits permet de représenter l'évolution des faits au cours du temps [26]. Bien que reposant sur la représentation d'un état de fait à un instant, cette méthode permet de représenter des entités perdurantes en considérant les entités représentées dans chacun des graphes successifs comme autant de parties temporelles. Cette méthode est également utilisée

pour représenter l'évolution de l'état des connaissances sur un domaine, avec une notion de version de graphe [36].

La problématique principale des approches par séries de graphes est qu'elles requièrent un certain synchronisme dans l'évolution des faits. En effet, dans le cas où on décrit un grand nombre d'entités susceptibles d'évoluer, ce mécanisme a pour défaut de nécessiter la duplication de l'entière du graphe pour chaque modification individuelle, ce qui rend son usage lourd dans le cas où les entités évoluent fréquemment et indépendamment les unes des autres. [36] propose une approche pour corriger ce défaut. En mettant en avant que le triplet est un élément insécable de tout graphe RDF et que celui-ci ne peut être qu'ajouté ou supprimé (l'édition correspondant à une suppression suivie d'un ajout), leur approche se base sur la représentation des modifications subies d'un graphe à l'autre.

Dans le contexte de séries de graphes, une difficulté est l'identification des ressources décrites d'un graphe à l'autre. L'idée sous-jacente à ces séries de graphes est généralement qu'une entité connaissant une évolution change d'identité. Autrement dit, il n'y a pas de raison qu'une ressource décrite dans deux graphes successifs possède le même URI dans chacun des graphes. En pratique d'ailleurs, l'usage d'un unique URI peut être problématique car la combinaison de deux graphes peut conduire à des connaissances contradictoires (car les connaissances au sein des graphes ne sont pas temporalisées). Afin de pouvoir représenter l'équivalence ou la succession des ressources décrites, des approches telles que le Change Bridge du projet SAMPO [26] permettent de lier une à une les ressources des différents graphes, faisant apparaître explicitement la ligne temporelle d'une ressource. On trouve également des approches hybrides, combinant le modèle Fluenta et les Change Bridge [7].

4.3 Temporalité et formalisme de représentation

Du fait des enjeux que présente la représentation temporelle, de nombreux travaux se sont attachés à définir un cadre de représentation. La table 2 dresse une liste (non-exhaustive) de ces formalismes en comparant leurs possibilités de représentation (et non de raisonnement). A l'exception de TA-RDF [40] qui s'approche du modèle fluenta, l'ensemble des formalismes présentés utilisent une mécanique d'annotation de triplet. Nombre d'entre eux complètent leur formalisme d'une syntaxe propre accompagnée d'un algorithme permettant de traduire les connaissances ainsi représentées en RDF classique [6, 39, 45, 25]. Ces modèles sont par ailleurs souvent complétés par un langage de requête propre généralement dérivé de SPARQL [6, 40]. Concernant le formalisme retenu, on constate une dominance des approches par points et par intervalles mathématiques, qui ont pour avantage de faciliter l'ancrage dans une temporalité précise. Enfin, les enjeux de représentation tels que l'instant présent ou la représentation de périodes temporelles vagues sont intégrés de manière disparate au sein des diverses approches.

Approche	Marqueur Présent	Temporalité vague	Représentation du temps
stRDF[6]	✓		Intervalles mathématiques
Temporal RDF[21, 22]	✓	Temporalité Anonyme	Point
tRDF[39]	✓	Triplet indéterminé	Point
Annotated RDF[49]			Intervalles fermés + Opérateurs
aRDF[45]		Pondération possible	Variable, mais nécessite une relation d'ordre
[33]	✓		Point
[25]		Pondération (Markov)	Allen + Intervalles mathématiques
[19]	✓		Intervalles mathématiques
TA-RDF [40]			Point
[14]		Pondération	Intervalles ouverts

TABLE 2 – Comparaison de différents formalismes de représentation temporelle

5 Conclusion

Lors de la conception d'une ontologie intégrant une temporalisation des faits, il est nécessaire de s'interroger sur la légitimité de temporaliser tel ou tel fait selon la conception du temps retenue. Si certains travaux [42] suggèrent que l'éternalisme serait l'alternative la plus adaptée dans un contexte physique, d'autres [44] mettent en avant l'utilité du possibilisme dans le contexte des humanités numériques par exemple. Selon le cadre de représentation de l'ontologie développée, il est donc bon de s'interroger sur la représentation la plus adaptée. Ceci étant acté, une représentation des composantes temporelles doit être choisie en fonction du contexte. Dans celui des humanités numériques par exemple, il sera certainement nécessaire de retenir un formalisme considérant des composantes temporelles imprécises/incertaines [10]. Dans certaines circonstances, la temporalisation pourra s'effectuer en se concentrant sur la représentation exclusive d'évènements [27]. Enfin, il sera important de s'interroger sur la technique à utiliser afin de lier les faits et les composantes temporelles exprimées à l'aide du formalisme retenu. Dans cet article, nous fournissons des éléments qualitatifs permettant d'orienter le choix selon la nature des données représentées. Des travaux actuellement en cours visent à comparer quantitativement ces divers formalismes, notamment concernant l'impact du nombre de propriétés moyen liant les entités, la propension d'une entité à évoluer, le degré de corrélation entre le changement de plusieurs propriétés d'une même entité ainsi que la syn-

chronicité avec laquelle les différentes entités évoluent.

Références

- [1] James F Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11) :832–843, 1983.
- [2] James F Allen and Patrick J Hayes. A common-sense theory of time. In *IJCAI*, volume 85, pages 528–531. Citeseer, 1985.
- [3] James F Allen and Patrick J Hayes. Moments and points in an interval-based temporal logic. *Computational Intelligence*, 5(3) :225–238, 1989.
- [4] Robert Arp, Barry Smith, and Andrew D Spear. *Building ontologies with basic formal ontology*. Mit Press, 2015.
- [5] Sotiris Batsakis and Euripides GM Petrakis. Sowl : a framework for handling spatio-temporal information in owl 2.0. In *Int. Workshop on Rules and Rule Markup Languages for the Semantic Web*, pages 242–249. Springer, 2011.
- [6] Konstantina Bereta, Panayiotis Smeros, and Manolis Koubarakis. Representation and querying of valid time of triples in linked geospatial data. In *The Semantic Web : Semantics and Big Data : 10th International Conference, ESWC 2013, Montpellier, France, May 26-30, 2013. Proceedings 10*, pages 259–274. Springer, 2013.
- [7] Camille Bernard, Marlène Villanova-Oliver, Jerome Gensel, and Hy Dao. Modeling changes in territorial partitions over time : Ontologies tsn and tsn-change. In *Proceedings of the 33rd Annual ACM Symposium on Applied Computing Pages (SAC '18)*, page 866–875, 04 2018.
- [8] Fabien Besnard. Time of philosophers, time of physicists, time of mathematicians. *arXiv preprint arXiv :1104.4551*, 2011.
- [9] Bertram C Bruce. A model for temporal references and its application in a question answering program. *Artificial intelligence*, 3 :1–25, 1972.
- [10] George Bruseker, Nicola Carboni, and Anaïs Guillem. Cultural heritage data management : the role of formal ontology and cidoc crm. *Heritage and Archaeology in the Digital Age*, pages 93–131, 2017.
- [11] Jeremy J Carroll, Christian Bizer, Pat Hayes, and Patrick Stickler. Named graphs. *Journal of Web Semantics*, 3(4) :247–267, 2005.
- [12] William Charles, Nathalie Hernandez, and Nathalie Aussenac-Gilles. Hht : An approach for representing temporally-evolving historical territories. In *Proc. of Extended Semantic Web Conference 2023, Hersonissos, Gr.*, page to Appear, 2023.
- [13] Simon JD Cox and SM Richard. A geologic timescale ontology and service. *Earth Science Informatics*, 8 :5–19, 2015.
- [14] Maximilian Dylla, Mauro Sozio, and Martin Theobald. Resolving temporal conflicts in inconsistent rdf knowledge bases. *Datenbanksysteme für Business, Technologie und Web (BTW)*, 2011.
- [15] Amer Ezoji and Nada Matta. Territorial knowledge ontology as a guide for the identification of resource of the territory toward sustainability. In *Proceedings of the Design Society : International Conference on Engineering Design*, volume 1, pages 3391–3400. Cambridge University Press, 2019.
- [16] Aldo Gangemi, Nicola Guarino, Claudio Masolo, Alessandro Oltramari, and Luc Schneider. Sweetening ontologies with dolce. In *Knowledge Engineering and Knowledge Management : Ontologies and the Semantic Web : Proc. of 13th International Conference, EKAW 2002 Sigüenza, Spain, Oct. 1–4, 2002*, pages 166–181. Springer, 2002.
- [17] José Miguel Giménez García. *Formalizing, Capturing, and Managing the Context of Statements in the Semantic Web*. Theses, Université de Lyon, July 2022.
- [18] Patrick Golden and Ryan Shaw. Nanopublication beyond the sciences : the periodo period gazetteer. *PeerJ Computer Science*, 2 :e44, 2016.
- [19] Fabio Grandi. Multi-temporal rdf ontology versioning. In *IWOD@ ISWC*, 2009.
- [20] Pierre Grenon and Barry Smith. Snap and span : Towards dynamic spatial ontology. *Spatial Cognition & Computation*, 4 :104 – 69, 2004.
- [21] Claudio Gutierrez, Carlos Hurtado, and Alejandro Vaisman. Temporal rdf. In *The Semantic Web : Research and Applications : Second European Semantic Web Conference, ESWC 2005, Heraklion, Crete, Greece, May 29–June 1, 2005. Proceedings 2*, pages 93–107. Springer, 2005.
- [22] Claudio Gutierrez, Carlos A Hurtado, and Alejandro Vaisman. Introducing time into rdf. *IEEE Transactions on Knowledge and Data Engineering*, 19(2) :207–218, 2006.
- [23] Olaf Hartig. Foundations of rdf* and sparql* : (an alternative approach to statement-level metadata in rdf). In Juan Reutter and Divesh Srivastava, editors, *AMW 2017 11th Alberto Mendelzon International Workshop on Foundations of Data Management and the Web, Montevideo, Uruguay*, volume 1912, 2017.
- [24] Pascal Hitzler. A review of the semantic web field. *Communications of the ACM*, 64(2) :76–83, 2021.
- [25] Jakob Huber. Temporal reasoning for rdf (s) : A markov logic based approach. Technical report, University of Mannheim, G., 2014.
- [26] Tomi Kauppinen and Eero Hyvönen. Modeling and reasoning about changes in ontology time series. *Ontologies : A handbook of principles, concepts and applications in information systems*, pages 319–338, 2007.

- [27] Gerard Kuys and Ansgar Scherp. Representing persons and objects in complex historical events using the event model f. *Journal of Open Humanities Data*, 8, 2022.
- [28] Maxime Lefrançois, Jarmo Kalaoja, Takoua Ghariani, and Antoine Zimmermann. The SEAS Knowledge Model. Research Report Deliverable 2.2, ITEA2 12004 Smart Energy Aware Systems, January 2017.
- [29] Jixin Ma. Ontological considerations of time, meta-predicates and temporal propositions. *Applied Ontology*, 2(1) :37–66, 2007.
- [30] Frank Manola, Eric Miller, Brian McBride, et al. Rdf primer. *W3C recommendation*, 10(1-107) :6, 2004.
- [31] Drew McDermott. A temporal logic for reasoning about processes and plans. *Cognitive science*, 6(2) :101–155, 1982.
- [32] Neil McKinnon. The endurance/perdurant distinction. *Australasian Journal of Philosophy*, 80(3) :288–306, 2002.
- [33] Boris Motik. Representing and querying validity time in rdf and owl : A logic-based approach. *Journal of Web Semantics*, 12 :3–21, 2012.
- [34] Vinh Nguyen, Olivier Bodenreider, and Amit Sheth. Don't like rdf reification? making statements about statements using singleton property. In *Proceedings of the 23rd international conference on World wide web*, pages 759–770, 2014.
- [35] Vinh Nguyen, Olivier Bodenreider, Krishnaprasad Thirunarayan, Gang Fu, Evan Bolton, N ria Queralt Rosinach, Laura I Furlong, Michel Dumontier, and Amit Sheth. On reasoning with rdf statements about statements using singleton property triples. *arXiv preprint arXiv :1509.04513*, 2015.
- [36] Damyan Ognyanov and Atanas Kiryakov. Tracking changes in rdf(s) repositories. In *Knowledge Engineering and Knowledge Management : Ontologies and the Semantic Web : 13th International Conference, EKAW 2002 Sig enza, Spain, October 1–4, 2002 Proceedings*, pages 373–378. Springer, 2002.
- [37] Daniel Peterson and Michael Silberstein. *Relativity of simultaneity and eternalism : In defense of the block universe*. Springer, 2010.
- [38] Axel Polleres, Aidan Hogan, Renaud Delbru, and J rgen Umbrich. Rdfs and owl reasoning for linked data. In *Reasoning Web. Semantic Technologies for Intelligent Data Access : 9th International Summer School 2013, Mannheim, G., July 30–Aug. 2, 2013. Proceedings*, pages 91–149. Springer, 2013.
- [39] Andrea Pugliese, Octavian Udrea, and VS Subrahmanian. Scaling rdf with time. In *Proceedings of the 17th international conference on World Wide Web*, pages 605–614, 2008.
- [40] Alejandro Rodriguez, Robert McGrath, Yong Liu, James Myers, and I Urbana-Champaign. Semantic management of streaming data. In *Proc. of the 2nd International Workshop on Semantic Sensor Networks (SSN09), collocated with the 8th Int. Semantic Web Conference (ISWC-2009), Washington DC, USA*, volume 522 of *CEUR Workshop Proceedings*, pages 80–95, 2009.
- [41] Anisa Rula, Matteo Palmonari, Andreas Harth, Stefan Stadtm ller, and Andrea Maurino. On the diversity and availability of temporal information in linked open data. In *The Semantic Web–ISWC 2012 : 11th International Semantic Web Conference, Boston, MA, USA, November 11–15, 2012, Proceedings, Part I 11*, pages 492–507. Springer, 2012.
- [42] Simon Saunders. How relativity contradicts presentism. *Royal Institute of Philosophy Supplements*, 50 :277–292, 2002.
- [43] Tom Stoneham. Time and truth : The presentism-eternalism debate. *Philosophy*, 84(2) :201–218, 2009.
- [44] Fumiaki Toyoshima. Ontology of time for the digital humanities : A foundational view. In *Proceedings of the Joint Ontology Workshop JOWO 2019*, volume 2518 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2019.
- [45] Octavian Udrea, Diego Reforgiato Recupero, and VS Subrahmanian. Annotated rdf. *ACM Transactions on Computational Logic (TOCL)*, 11(2) :1–41, 2010.
- [46] Hsien-Tseng Wang and Abdullah Uz Tansel. Temporal extensions to rdf. *Journal of Web Engineering*, 2019.
- [47] Christopher A. Welty and Richard Fikes. A reusable ontology for fluents in OWL. In Brandon Bennett and Christiane Fellbaum, editors, *Formal Ontology in Information Systems, Proceedings of the Fourth International Conference, FOIS 2006, Baltimore, Maryland, USA*, volume 150 of *Frontiers in Artificial Intelligence and Applications*, pages 226–236. IOS Press, 2006.
- [48] Fu Zhang, Ke Wang, Zhiyin Li, and Jingwei Cheng. Temporal data representation and querying based on rdf. *IEEE Access*, 7 :85000–85023, 2019.
- [49] Antoine Zimmermann, Nuno Lopes, Axel Polleres, and Umberto Straccia. A general framework for representing, reasoning and querying with annotated semantic web data. *Journal of Web Semantics*, 11 :72–95, 2012.