

# Rappresentazione, costruzione e visualizzazione di risorse terminologiche diacroniche nell'era del web semantico

Silvia Piccini, Matteo Abrate, Andrea Bellandi, Emiliano Giovannetti

## 1. Introduzione

Il presente contributo ha l'obiettivo di illustrare *DIATERM*, un modello creato per rappresentare formalmente l'evoluzione diacronica di concetti e termini in un dato dominio, nel quadro del web semantico.

Al di là di poche eccezioni tutte concentrate nell'ultimo ventennio inter al. (Van Campenhoudt 1997; Humbley 2011<sup>1</sup>; Zanola 2014; Piccini 2016)<sup>2</sup>, la dimensione diacronica è stata a lungo trascurata in terminologia, a causa dell'ap-proccio fondamentalmente onomasiologico, normativo e statico alla base della teoria 'classica' della terminologia.

Inaugurata da Wüster nel 1968, tale visione tradizionale si fonda infatti su una semiotica di stampo neopositivista (Slodzian 2000) secondo la quale il termine – entità fondamentalmente monosemica – instaurerebbe con il concetto – entità

<sup>1</sup> Cfr. i riferimenti bibliografici del presente contributo.

<sup>2</sup> La bibliografia ivi riportata non ha la pretesa di essere esaustiva, ma solo indicativa. Vivo è infatti l'interesse registrato negli ultimi anni verso la variazione diacronica in terminologia.

Silvia Piccini, CNR, National Research Council of Italy, Italy, [silvia.piccini@ilc.cnr.it](mailto:silvia.piccini@ilc.cnr.it), 0000-0002-2584-0191  
Matteo Abrate, CNR, National Research Council of Italy, Italy, [matteo.abrate@iit.cnr.it](mailto:matteo.abrate@iit.cnr.it),  
Andrea Bellandi, CNR, National Research Council of Italy, Italy, [andrea.bellandi@ilc.cnr.it](mailto:andrea.bellandi@ilc.cnr.it), 0000-0002-1900-5616  
Emiliano Giovannetti, CNR, National Research Council of Italy, Italy, [emiliano.giovannetti@ilc.cnr.it](mailto:emiliano.giovannetti@ilc.cnr.it), 0000-0002-0716-1160

FUP Best Practice in Scholarly Publishing (DOI 10.36253/fup\_best\_practice)

Silvia Piccini, Matteo Abrate, Andrea Bellandi, Emiliano Giovannetti, *Rappresentazione, costruzione e visualizzazione di risorse terminologiche diacroniche nell'era del web semantico*, pp. 125-139, © 2021 Author(s), CC BY 4.0 International, DOI 10.36253/978-88-5518-364-2.12, in Claudio Grimaldi, Maria Teresa Zanola (edited by), *Terminologie e vocabolari. Lessici specialistici e tesauri, glossari e dizionari*, © 2021 Author(s), content CC BY 4.0 International, metadata CC0 1.0 Universal, published by Firenze University Press ([www.fupress.com](http://www.fupress.com)), ISSN 2704-5846 (online), ISBN 978-88-5518-364-2 (PDF), DOI 10.36253/978-88-5518-364-2

monodimensionale – una corrispondenza univoca perfetta, indifferente a qualsiasi variazione culturale e temporale. Secondo una visione già aristotelica, i termini, quindi, rifletterebbero una realtà unica, immobile ed atemporale, aderendo perfettamente a concetti e categorie di ordine logico e universale.

In questo rapporto ideale con il concetto, il termine perde il suo valore di segno linguistico per divenire una semplice etichetta, un simbolo e la terminologia si tramuta in una nomenclatura, in un sistema di denominazione.

Tale concezione classica nasce dall'esigenza di garantire una comunicazione priva di ambiguità tra gli studiosi, necessità questa che non può essere soddisfatta senza politiche di standardizzazione e normalizzazione. E di fatto, in una fase di «scienza normale», quando cioè la dottrina aderisce ad un «paradigma» stabilito e condiviso (Kuhn 1970), il sistema di termini è standard, esplicitamente definito e utilizzato consensualmente all'interno di una comunità scientifica. Tuttavia, quando si assiste a rivoluzioni epistemologiche, che rimettono in discussione le assunzioni basilari all'interno della teoria scientifica dominante, il sistema terminologico può subire cambiamenti, a volte anche profondi.

È in questi contesti che le testimonianze scritte acquisiscono un valore fondamentale per il terminologo. In una prospettiva storico-diacronica il testo, inteso qui in senso lato come espressione di una attività intellettuale organizzata in modo più o meno coerente, costituisce l'unica e imprescindibile fonte per ricostruire e comprendere la relazione profonda che intercorre tra teorizzazione e terminologia.

Possiamo pertanto affermare che la terminologia storico-diacronica è per definizione una «terminologia testuale» (Bourigault, Slodzian 1999), una terminologia del *discours*, della *parole*. La dimensione storico-diacronica pone al centro la *Ist-Norm*, l'uso reale della parola in un (con)testo, restituendo così al termine la sua vera natura di segno linguistico, nella complessità dei rapporti che esso stabilisce con gli altri termini in sincronia o nel divenire.

Molte sono le fonti testuali che hanno costituito e continuano a costituire una miniera preziosa di informazioni per lo studioso della storia della terminologia.

Nel presente contributo gli esempi proposti nei vari paragrafi sono tratti dall'opera *In Sphaeram Ioannis Sacro Bosco Commentarius* realizzata dall'astronomo e matematico gesuita Cristoforo Clavio (1538-1612) e dalle pagine manoscritte del padre della linguistica generale Ferdinand de Saussure (1857-1913).

Si tratta di due testualità diverse – l'una chiusa e definita, propria del testo a stampa e l'altra incompleta ed aperta, caratteristica delle pagine manoscritte – che permettono al terminologo di studiare due momenti differenti di interazione intensa tra teorizzazione e terminologia.

Il Commentario al *Tractatus de Sphaera*, infatti, nel ripercorrere la storia della disciplina astronomica, consente al terminologo di ricostruire gli effetti prodotti sul sistema terminologico dall'onda d'urto generata dalla rivoluzione scientifica che si ebbe a cavallo tra il XV ed il XVI secolo ad opera di grandi scienziati quali Niccolò Copernico, Tycho Brahe, Galileo Galilei, Giovanni Keplero.

Le pagine manoscritte del linguista ginevrino Ferdinand de Saussure (1857-1913), al contrario, consentono allo studioso di osservare *in vivo* il delinearci di

una terminologia nel momento in cui viene fondata una nuova scienza e si definiscono i principi che ne andranno a costituire il paradigma.

Il contributo si articola in sei paragrafi: dopo una breve introduzione sulla rappresentazione di aspetti diacronici nell'ambito del web semantico (§2), verrà descritto in dettaglio il modello *DIATERM* (§3). Quindi verrà presentato LexO, uno strumento volto a costruire ed interrogare risorse terminologiche diacroniche (§4). Nel paragrafo 5 verranno proposti dei prototipi di visualizzazioni interattive utili ad evidenziare l'evoluzione sia sul piano terminologico sia sul piano concettuale. Il contributo si concluderà con considerazioni sulle future linee di ricerca (§6).

## 2. La rappresentazione dell'evoluzione dell'informazione nel web semantico

L'imperativo metodologico sul quale si fonda il presente lavoro è stato quello di costruire un modello che consentisse di creare risorse terminologiche diacroniche «FAIR», ossia facilmente reperibili, accessibili, interoperabili e riutilizzabili all'interno di una comunità scientifica (Wilkinson *et al.* 2016).

Di qui la scelta di adottare le tecnologie proprie del web semantico (RDF, OWL, SPARQL e SWRL) ed il paradigma dei *Linked Data*, seguendo una tendenza ormai sempre più diffusa tra i lessicografi ed i terminologi computazionali.

Rappresentare l'informazione temporale con le tecnologie del web semantico costituisce, tuttavia, un'operazione complessa.

RDF (*Resource Description Framework*) – il modello standard di codifica dei dati – e OWL (*Ontology Web Language*) – il linguaggio standard per la rappresentazione e la condivisione delle ontologie sul web – si fondano, infatti, su una struttura delle asserzioni (*statements*) in forma di triple composte da soggetto-predicato-oggetto.

Essi sottintendono, quindi, logiche basate su predicati binari, incompatibili con la necessità di associare una validità temporale ad una certa relazione che lega due entità. In tal caso, infatti, la relazione si trasforma da binaria in ternaria.

In letteratura sono state proposte diverse soluzioni per la rappresentazione di informazione dinamica nell'ambito del web semantico, quali ad esempio il *versioning* (Klein, Fensel 2001), la reificazione (Manola, Miller 2004), l'approccio a perduranti, anche detto dei *4D fluents* (Welty, Fikes 2006).

In *DIATERM* la scelta è stata orientata verso l'adozione di un approccio basato sulla reificazione delle relazioni N-arie.

Rispetto al modello dei fluenti, infatti, l'approccio N-ario comporta l'introduzione di un numero minore di entità (concetti e relazioni) e di asserzioni (assiomi) e garantisce così una maggior efficacia nei meccanismi di *reasoning* (Batsakis *et al.* 2017).

Più specificamente, in *DIATERM* le relazioni di arietà superiore a due sono espresse attraverso l'*ontology design pattern* della reificazione (Noy *et al.* 2006), metodo altamente raccomandato dal *Semantic Web Best Practices and Deployment Working Group*.

Reificare significa trasformare un predicato in oggetto, ovvero rappresentare una relazione come classe. Come è noto, infatti, non sempre è possibile scom-

porre una relazione con arietà maggiore di due in una concatenazione di relazioni binarie secondo l'approccio *daisy-chaining*, senza che il significato originario venga compromesso<sup>3</sup>.

Così ad esempio l'asserzione «il termine *t* è caratterizzato dal senso *s* che denota il concetto *c*» può essere considerata equivalente al concatenamento delle due asserzioni seguenti «*t* è caratterizzato dal senso *s*» e «*s* denota il concetto *c*». Al contrario, invece, la relazione implicata nell'asserzione «il termine *t* è caratterizzato dal senso *s* nell'intervallo di tempo *I*», è intrinsecamente ternaria, poiché lega indissolubilmente le tre entità coinvolte e non può essere scomposta in relazioni binarie più elementari.

La reificazione offre in questi casi una soluzione, sebbene lo scotto da pagare sia alto in termini di complessità della formalizzazione e di tempi di computabilità.

### 3. Il modello *DIATERM*

Il modello *DIATERM* (Figura 1) presenta un'architettura articolata su tre livelli.

L'asse diacronico è trasversale a tutti e tre i piani, in quanto il mutamento può verificarsi a livello testuale, terminologico e concettuale. In altri termini, può essere il testo che – nella sua fase di gestazione, prima di essere dato alla stampa – viene sottoposto a riscritture continue, come il caso di Saussure mette bene in mostra; o ancora possono essere il piano terminologico e/o il piano concettuale a subire variazioni. In alcuni casi il mutamento investe tutti e tre i piani.

La separazione del piano linguistico e del piano concettuale, ovvero la distinzione tra significato e concetto – questione teoricamente abbastanza dibattuta – si rivela molto utile dal punto di vista metodologico, in quanto permette di far emergere chiaramente se il cambiamento di denominazione corrisponda di fatto ad un mutamento significativo nella teoria.

Prima di entrare nel dettaglio del modello, una considerazione di carattere generale è necessaria.

In accordo con la filosofia dei *Linked Data* e per contrastare l'eterogeneità del web semantico, in *DIATERM* sono stati integrati, laddove possibile, vocabolari diversi. Così ad esempio le entità precedute dal prefisso “*dcmitype*” sono state mutate dal vocabolario della *Dublin Core Metadata Initiative* (DCMI)<sup>4</sup>, uno dei più noti sistemi standard di metadatozione delle risorse in formato elettronico. Il prefisso “*time*” indica, invece, entità che appartengono ad *OWL-Time*<sup>5</sup>, un'ontologia realizzata dal W3C per descrivere il contenuto temporale delle pagine web, come ad esempio date, relazioni topologiche tra istanti e intervalli, durata di intervalli. Altri vocabolari utilizzati verranno indicati in nota.

<sup>3</sup> Cfr. Grewe 2010.

<sup>4</sup> <<https://www.dublincore.org/>> (2020-04-29).

<sup>5</sup> <<https://www.w3.org/TR/owl-time/>> (2020-04-29).

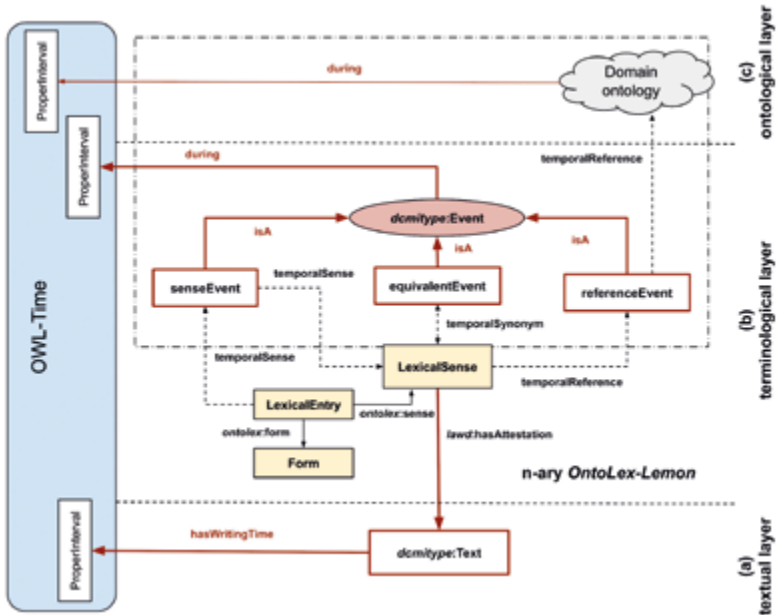


Figura 1 – Il modello DIATERM

### 3.1 Il livello ontologico in DIATERM

Nel *livello ontologico* di DIATERM vengono strutturati e descritti formalmente, attraverso una ontologia in OWL, i concetti denotati dai termini e la loro evoluzione.

Poiché come abbiamo sottolineato nel paragrafo 2 la reificazione costituisce il meccanismo alla base della rappresentazione temporale dell'informazione, la classe `dcmltype:Event`<sup>6</sup> gioca un ruolo centrale in DIATERM. Essa rappresenta genericamente la reificazione di una proprietà che sussiste in un dato intervallo di tempo *t*.

Accanto alla classe `dcmltype:Event`, sono state introdotte due proprietà:

- la proprietà `during`, il cui dominio e codominio sono rappresentati rispettivamente dalla classe `dcmltype:Event` e dalla classe `time:ProperInterval`<sup>7</sup>. La proprietà `during` è stata modellata come sotto-proprietà del tipo `dcterms:date`<sup>8</sup>, anch'esso tratto dal vocabolario DCMI;
- la relazione `temporalProperty`, che collega le istanze della classe `dcmltype:Event` rispettivamente agli individui *source* e *target* della relazione reificata. Come vedremo, le relazioni convertite in proprietà temporali sono state formalizzate come sotto-proprietà di `temporalProperty`.

<sup>6</sup> <<http://purl.org/dc/dcmtype/Event>> (2020-04-29).

<sup>7</sup> <<https://www.w3.org/TR/owl-time/#time:ProperInterval>> (2021-06-29)

<sup>8</sup> <<http://purl.org/dc/elements/1.1/date>> (2021-06-29)

### 3.2 Il livello terminologico in *DIATERM*

Passiamo ora al *livello terminologico*. Qui le informazioni linguistiche dei termini vengono formalizzate secondo il modello *OntoLex-Lemon* (McCrae *et al.* 2017), standard *de facto* per la rappresentazione della componente lessicale di ontologie nel web semantico<sup>9</sup>. L'adozione di tale modello è stata dettata principalmente da due motivi. Anzitutto, *OntoLex-Lemon* consente di creare lessici in RDF, modello dati fondante dell'architettura del web semantico. In secondo luogo, il nucleo di *OntoLex-Lemon* è caratterizzato da una distinzione, che abbiamo detto essere centrale nel presente lavoro, tra il piano linguistico dei significati, istanziati nella classe `ontolex:LexicalSense`, ed il piano ontologico dei concetti, che ricevono, invece, una descrizione formale in un'ontologia al di fuori del modello.

Nel livello terminologico di *DIATERM* ritroviamo, quindi, le tre classi che costituiscono il cuore del modello *OntoLex-Lemon*: la classe `ontolex:LexicalEntry` le cui istanze sono costituite da entrate lessicali, siano esse realizzate come parole semplici (`ontolex:Word`), polirematiche (`ontolex:MultiwordExpression`) o affissi (`ontolex:Affix`); la classe `ontolex:LexicalForm`, le cui istanze sono rappresentate dalle differenti realizzazioni morfologiche di un'entrata lessicale; la classe `ontolex:LexicalSense`, i cui individui rappresentano i vari significati che il lessema può assumere nei diversi contesti.

Tutte le relazioni previste dal modello *OntoLex-Lemon* sono state reificate e convertite in sottoclassi della classe `dcmltype:Event`. Si hanno quindi la classe `senseEvent`, reificazione della relazione `ontolex:sense` che collega un individuo della classe `ontolex:LexicalEntry` a un significato specifico; la classe `referenceEvent`, reificazione della relazione `ontolex:reference` che lega il significato di un'entrata lessicale ad un concetto dell'ontologia; le classi `equivalentEvent`, `incompatibleEvent`, `broaderEvent` e `narrowerEvent`, che rappresentano rispettivamente la reificazione delle relazioni `lexinfo:synonym`, `lexinfo:antonym`, `lexinfo:hypernym` e `lexinfo:hyponym`.

La classe `dcmltype:Event` – e le sue sottoclassi – sono infine collegate alla classe `time:ProperInterval` tramite la proprietà `during`, che indica l'intervallo di tempo in cui si verifica l'evento in questione.

Infine, come sotto-proprietà di `temporalProperty` sono state introdotte delle nuove relazioni, quali `temporalSynonym`, `temporalAntonym`, `temporalHypernym` e `temporalHyponym`, che legano gli individui della classe `ontolex:LexicalSense` alle istanze delle classi risultanti dalla reificazione sopra citate.

Supponiamo di voler rappresentare formalmente il fatto che un termine  $t_1$  sia sinonimo di un termine  $t_2$  in un determinato intervallo di tempo  $I$ . L'istanza  $t_1$  della classe `ontolex:LexicalSense` viene collegata attraverso la relazione `temporalSynonym` ad un individuo della classe `equivalentEvent`. Quest'ultimo, a sua volta, viene legato attraverso la stessa relazione `temporalSynonym` al termine  $t_2$  appartenente

<sup>9</sup> Occorre sottolineare che *OntoLex-Lemon* non fornisce un vocabolario per la descrizione delle categorie linguistiche, ma raccomanda a tal fine l'utilizzo di LexInfo. In questo lavoro le categorie linguistiche sono mutate da tale vocabolario <<https://www.lexinfo.net/ontology/2.0/lexinfo>> (2020-04-29).

alla classe `ontolex:LexicalSense` e tramite la relazione `during` ad un intervallo di tempo preciso, che definisce la validità dell'evento di sinonimia.

Come si evince da questo esempio, il meccanismo di reificazione introduce una maggiore complessità nella formalizzazione: la relazione di sinonimia non lega più direttamente i due termini  $t_1$  e  $t_2$ , ma viene spezzata dall'introduzione di una nuova classe, nell'esempio sopra riportato `equivalentEvent`, che ne modifica il dominio ed il codominio, rendendo così più complessi i meccanismi di *reasoning*.

### 3.3 Il livello testuale in *DIATERM*

Nel *livello testuale* vengono dettagliate, attraverso lo schema di metadateazione standard *DublinCore*<sup>10</sup>, le informazioni principali (autore, titolo, catalogazione archivistica, ecc.) concernenti le fonti testuali dalle quali sono estratti i termini.

Il concetto di attestazione svolge in *DIATERM* un ruolo centrale, in quanto, come abbiamo già sottolineato, la terminologia storico-diacronica è per definizione una terminologia testuale: è la pagina scritta che consente al terminologo di ripercorrere le vicende di un termine.

Le entità che costituiscono il livello testuale sono le seguenti:

- la classe `dcmltype:Text`<sup>11</sup>, le cui istanze sono rappresentate da entità molto diverse tra loro, come i libri, gli articoli pubblicati, gli appunti, le bozze di opere incomplete, ecc. Il concetto di testo deve essere, infatti, inteso in senso lato come un «enunciato complesso, orale o scritto, considerato un'entità unitaria in base a proprietà particolari quali la compattezza morfosintattica e l'unità di significato»<sup>12</sup>;
- la proprietà `lawd:hasAttestation`<sup>13</sup>, che collega un significato, cioè un'istanza della classe `ontolex:LexicalSense`, ad un'istanza della classe `dcmltype:Text`;
- la proprietà `hasWritingTime` che, definita come sotto-proprietà di `dcterms:date`, collega un individuo della classe `dcmltype:Text` ad un individuo appartenente alla classe `time:ProperInterval`. Per quanto concerne i testi pubblicati, è stata mutuata dal vocabolario della DCMI la relazione `dcterms:issued`<sup>14</sup>. In tal caso, viene considerato come fine del periodo di scrittura l'anno di pubblicazione dell'opera.

L'adozione del linguaggio OWL nella formalizzazione dell'evoluzione diacronica di termini e concetti offre grandi potenzialità, come ad esempio la possibilità di attribuire automaticamente certe informazioni temporali a partire dalla conoscenza già esplicitata nell'ontologia.

<sup>10</sup> Lo studioso può adottare anche altri sistemi di metadateazione (BIBFRAME, FRBR, ecc.), qualora li ritenga più opportuni ai fini della propria ricerca, senza che questo comprometta la validità del modello *DIATERM*.

<sup>11</sup> `<http://purl.org/dc/dcmitype/Text>` (2021-06-29).

<sup>12</sup> La definizione è tratta dal *Dizionario della lingua italiana* curato da Tullio De Mauro.

<sup>13</sup> La relazione è stata mutuata dall'ontologia LAWD (*Linked Ancient World Data*): `<http://lawd.info/ontology/hasAttestation>` (2020-04-29).

<sup>14</sup> `<http://purl.org/dc/terms/issued>` (2020-04-29).

A tal fine sono state elaborate delle regole SWRL che, sfruttando le relazioni stabilite da Allen, consentono di definire la validità temporale delle relazioni che legano termini e concetti.

Così ad esempio, a partire dalla datazione del testo, può essere automaticamente stabilito il periodo di attestazione di un senso di un termine, come mostra la regola seguente:

$$\begin{aligned} & \text{Text}(?t) \wedge \text{time:ProperInterval}(?i) \wedge \text{hasWritingTime}(?t, ?i) \wedge \\ & \text{ontolex:LexicalEntry}(?l) \wedge \text{ontolex:LexicalSense}(?s) \wedge \text{isAttestedIn}(?s, ?t) \wedge \\ & \text{ontolex:sense}(?l, ?s) \wedge \text{swrlx:createOWLThing}(?se, ?s) \rightarrow \\ & \text{SenseEvent}(?se) \wedge \text{temporalSense}(?l, ?se) \wedge \text{temporalSense}(?se, ?s) \wedge \text{during}(?se, ?i) \end{aligned}$$

L'utilizzo di tali regole SWRL agevola l'attività del terminologo nel popolamento della risorsa terminologica, evidenziando inoltre eventuali inconsistenze nella formalizzazione.

#### 4. La costruzione e l'interrogazione di una risorsa terminologica diacronica: LexO

Come si evince da quanto sopra esposto, l'adozione dell'*ontology design pattern* della reificazione rende l'attività di codifica dell'informazione temporale da parte del terminologo estremamente complessa e prona ad errori. L'utilizzo di editor ontologici come Protégé non aiuta il terminologo in questo processo di rappresentazione dell'informazione. Lo studioso si ritrova a dover trasformare manualmente relazioni in classi perdendosi così nella verbosità del modello, per recuperare con difficoltà la trasparenza delle informazioni prima della reificazione.

Al fine di consentire agli studiosi di creare risorse diacroniche astraendo dalla complessità degli aspetti di formalizzazione, è in corso di realizzazione una estensione di LexO (Bellandi *et al.* 2018 e Bellandi, 2021), un editor web, collaborativo e personalizzabile di lessici computazionali sviluppato presso l'Istituto di Linguistica Computazionale del CNR. I terminologi, grazie ad un'interfaccia semplice e intuitiva, potranno così costruire risorse diacroniche secondo gli standard del web semantico, senza tuttavia dover essere esperti dei formalismi sottiacenti.

Nella Figura 2 viene illustrata l'interfaccia grafica principale di LexO. Essa rispecchia la tripartizione lemma-forma-senso costitutiva del modello *OntoLex-Lemon* adottato in *DIATERM*. Nella parte sinistra dell'interfaccia sono elencate le entrate lessicali che costituiscono il dizionario, mentre sulla destra viene fornita la descrizione morfologica e semantica dell'entrata lessicale selezionata. A titolo esemplificativo, nella Figura 2 si riporta la descrizione del termine *articulus* tratto dal dizionario della terminologia saussuriana (Ruimy *et al.* 2013). Il riquadro rosso è deputato alla definizione delle informazioni principali concernenti il lemma, quali la categoria grammaticale, il genere ed il numero; nel riquadro azzurro vengono riportate e descritte le forme che eventualmente occorrono nel testo (nell'esempio *articuli*, la flessione plurale del termine). Il riquadro giallo è dedicato alla descrizione del senso, del quale viene data una definizione in linguaggio naturale e vengono formalizzate le eventuali relazioni intrattenute con i sensi degli altri termini del dizionario. In tal caso, poiché *articulus* costituisce



un termine polisemico, sono dati due riquadri gialli corrispondenti ciascuno ad uno dei due sensi con i quali il termine è usato negli scritti di Saussure.

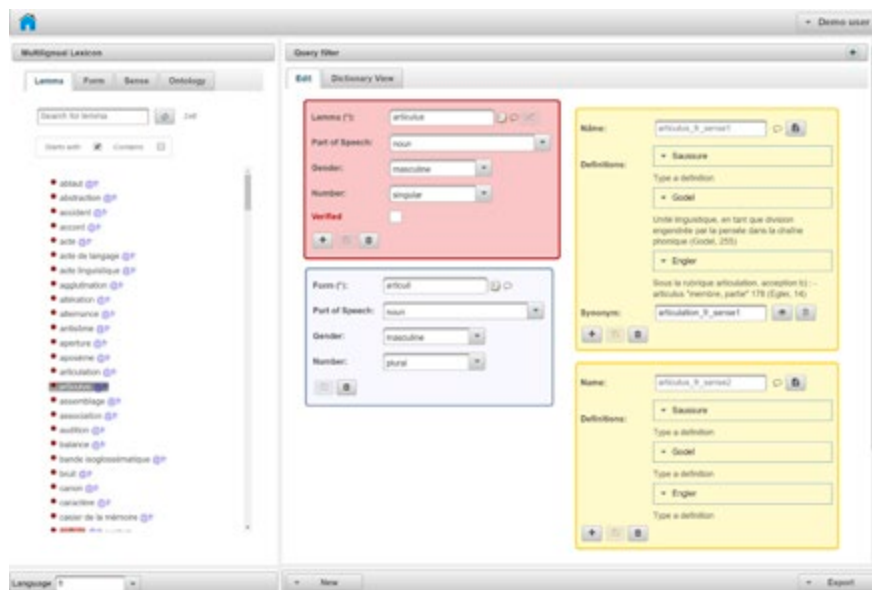


Figura 2 – Interfaccia principale di LexO: il termine saussuriano *articulus*

LexO permette anche di collegare il senso lessicale ad un concetto ontologico descritto in una ontologia di dominio esterna al lessico.

Allo stato attuale, il modello *DIATERM* non è stato implementato in LexO e lo strumento, pertanto, non offre la possibilità di editare direttamente dall'interfaccia la validità temporale dei sensi e delle loro relazioni. Tuttavia, esso può essere utilizzato per visualizzare ed interrogare una risorsa diacronica costruita secondo il modello N-ario.

Uno dei vantaggi principali di questa formalizzazione così complessa risiede, infatti, nella possibilità di effettuare ricerche sofisticate e complesse.

LexO consente due tipi di interrogazioni, che l'utente può formulare facilmente in linguaggio naturale grazie ad un'interfaccia *Question-Answering*, costituita da template predefiniti.

Il primo tipo di interrogazione permette allo studioso di conoscere l'evoluzione diacronica delle relazioni lessico-semantiche che i termini instaurano tra di loro. Supponiamo, ad esempio, che lo studioso di Saussure voglia sapere quali termini nel corso del tempo sono stati usati come sinonimi (o quasi-sinonimi) del termine *signe*, nella sua accezione di «entità psichica costituita da un significato ed un significante» (*signe\_sense3*). Il risultato della ricerca è illustrato sotto forma di grafico nella Figura 3.

Il lavoro costante, quasi ossessivo, condotto dal linguista ginevrino alla ricerca della denominazione più adatta, emerge immediatamente.

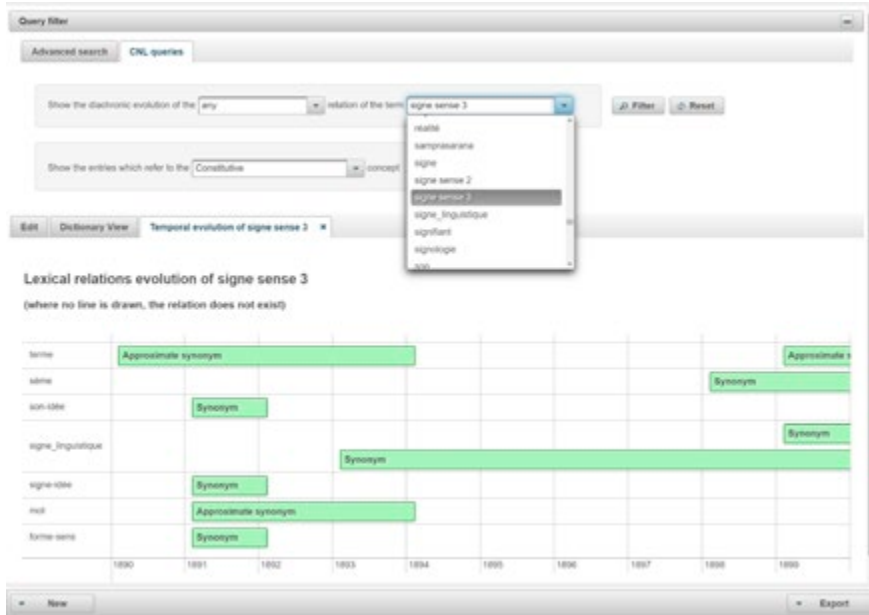


Figura 3 – I sinonimi del termine *signe* tra il 1890 e il 1900

Conscio, infatti, dell'ambiguità teorica generata dalla polisemia del termine *signe*<sup>15</sup>, Saussure prova ad introdurre nei suoi scritti diverse denominazioni. Alcune coesistono per periodi più o meno lunghi (*mot*, *terme*, *signe linguistique*); altre vengono quasi subito abbandonate, come ad esempio il termine *sème* confinato alle *Notes Item* (1899-1903) o le polirematiche *son-idée*, *signe-idée*, *forme-sense*, introdotte nella *Double Essence du langage* (1891-1892) nel tentativo di tradurre iconicamente nella lingua le riflessioni sulla natura diadica propria del segno linguistico. Anche questi neologismi saranno abbandonati e verrà conservato il termine di *signe*, usato ancora oggi in linguistica e semiotica.

LexO prevede anche un secondo tipo di interrogazione su base ontologica, che consente allo studioso di ottenere sotto forma di *synset* tutti i termini che sono stati utilizzati in sincronia e in diacronia per denotare un determinato concetto ontologico.

## 5. Rappresentazione visiva e confronto diacronico

Al fine di rendere ancora più esplicita allo studioso l'evoluzione di una risorsa diacronica, sono stati realizzati differenti prototipi di visualizzazione interattiva.

<sup>15</sup> Il termine *signe* è utilizzato dal maestro ginevrino per designare sia l'immagine acustica di cui si compone il segno linguistico sia il segno linguistico vero e proprio, ovvero la combinazione di immagine acustica e concetto.

Il primo esempio proposto è tratto dal già citato Commentario al *Tractatus de Sphaera* di Cristoforo Clavio. Attraverso una rappresentazione grafica ad albero (Figura 4), dove i concetti sono rappresentati in giallo, le istanze in viola ed i termini in marrone, viene illustrato il concetto di *primum mobile* in tre momenti teorici importanti. Il confronto tra queste tre rappresentazioni permette di ricostruire l'evolversi nel tempo del concetto e di osservarne le ripercussioni terminologiche.

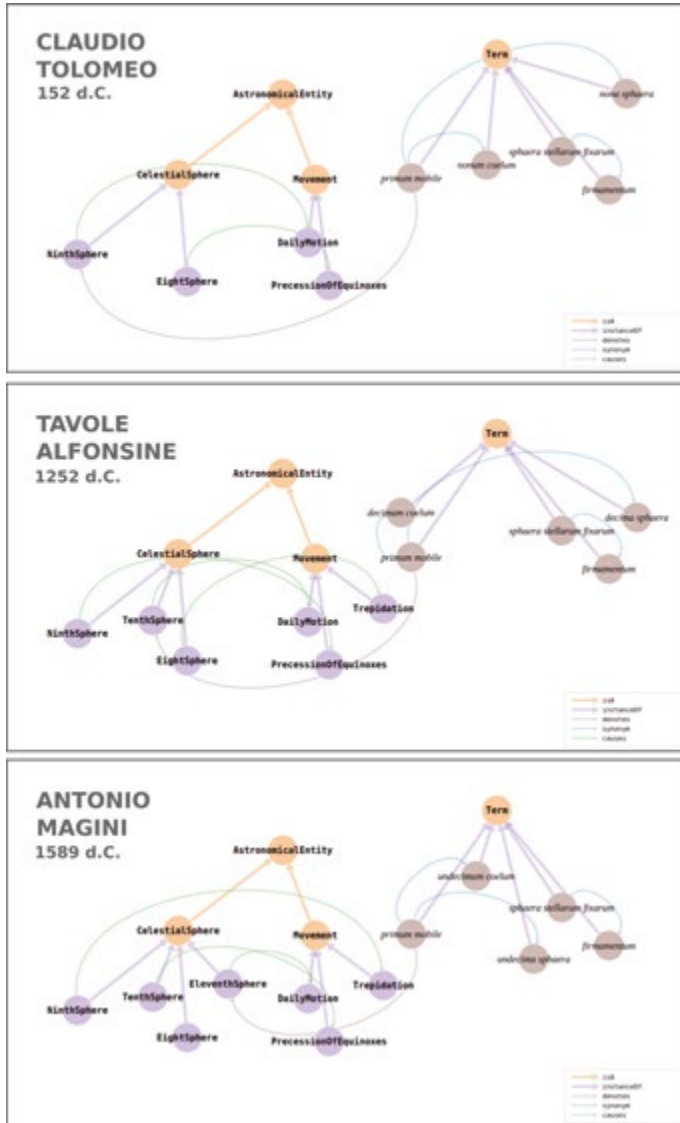


Figura 4 – L'evoluzione del concetto di *primum mobile* nel Commentario di Cristoforo Clavio

Come si vede dalla Figura 4, il termine *primum mobile* è stato utilizzato nel tempo per denotare sfere diverse. Mano a mano che nei secoli, infatti, si venivano a scoprire maggiori dettagli sul moto dei pianeti, diveniva necessario aggiungere delle sfere: così con Tolomeo il loro numero salì a nove (152 d.C., data dell'*Almagesto*), per dare spiegazione del movimento scoperto da Ipparco e noto con il nome di precessione degli equinozi; nelle Tavole alfonsine (1252) le sfere celesti divennero dieci per spiegare la causa del movimento di trepidazione scoperto dall'astronomo arabo Thābit ibn Qurra. Infine Magini, nel *Novae coelestium orbium theoricarum* (1589), teorizzò la necessità di una ulteriore undicesima sfera.

All'introduzione di una ulteriore sfera sul piano concettuale corrisponde, sul piano terminologico, un mutamento che si riflette anche nei rapporti di sinonimia tra i termini. Così ad esempio nelle pagine del Commentario di Clavio che illustrano la visione astronomica classica di impianto aristotelico, il termine *primum mobile* è usato come sinonimo di *firmamentum* e di *sphaera stellarum fixarum*; poche pagine dopo, quando è il sistema tolemaico ad essere descritto, *primum mobile* diviene, invece, sinonimo di *sphaera nona*.

Nonostante, come abbiamo visto, la rappresentazione concettuale e terminologica assuma un aspetto ad albero, la struttura dati su cui si fonda la visualizzazione non costituisce in realtà un albero *stricto sensu*, nell'accezione informatica del termine<sup>16</sup>.

Alcune tipologie di relazioni, infatti, rompono la gerarchia tassonomica padre-figlio, secondo la quale ogni nodo dell'albero può avere un solo arco uscente. Esistono, cioè, casi di ereditarietà multipla, per cui lo stesso nodo può presentare più archi uscenti, come ad esempio la relazione orientata *causes*. Inoltre, alcune tipologie di relazioni possono definire dei cicli, come nel caso della relazione transitiva di sinonimia.

Sebbene in tali casi si possa ricorrere ad una più generica struttura a grafo ed utilizzare appropriati algoritmi di layout, la scelta operata in questa sede è stata tuttavia quella di rimuovere temporaneamente le relazioni che creavano cicli e rappresentare l'insieme restante di nodi ed archi attraverso un *grafo orientato aciclico* (detto anche DAG, *Directed Acyclic Graph*). Questa struttura ammette relazioni ad ereditarietà multipla e presenta il vantaggio di mantenere l'aspetto gerarchico tipico dell'albero, una volta utilizzato un appropriato algoritmo di layout.

Così ad esempio le relazioni di sussunzione (*isA*) e istanziazione (*instanceOf*) vengono rappresentate nelle modalità tipiche di una struttura propriamente ad albero, dove i concetti più generali sono collocati nella parte alta della gerarchia concettuale mentre le istanze nella parte bassa. Una volta che questi nodi e archi sono stati tutti posizionati, si è proceduto alla reintroduzione delle relazioni che possono definire cicli.

Il grafo orientato aciclico risulta una struttura particolarmente atta a rappresentare l'evoluzione diacronica, in quanto consente di fondere strutture dati cor-

<sup>16</sup> Una struttura dati ad albero è costituita da nodi ed archi, che stabiliscono una relazione gerarchica tra i nodi. Per ogni nodo si può avere un solo arco uscente e più archi entranti.



do perdipiù l'onere dell'impiego di algoritmi iterativi di approssimazione della posizione ottimale dei nodi. Questo layout è stato introdotto da Bezerianos *et al.* (2010) per la rappresentazione di alberi genealogici, la cui struttura è affine a quella ad albero classica, pur presentando genitori multipli.

La tecnica di rappresentazione visiva è prossima a quella descritta nel primo esempio, ma si è scelto di lasciare visibili in trasparenza i nodi e gli archi relativi agli altri istanti di tempo.

Una tale visualizzazione consente allo studioso di avere una percezione immediata di come la terminologia si sia evoluta. Come già notato nell'esempio proposto nel paragrafo precedente, ad esempio, i composti introdotti da Saussure nel 1891-1892 per indicare il concetto di segno linguistico (*Relational\_Entity*) vengono abbandonati nel 1910-1911, data del terzo Corso di linguistica generale. In tale occasione Saussure introduce per la prima volta i termini *signifié* e *signifiant* per indicare rispettivamente la parte concettuale e la parte 'materiale' del segno linguistico. Il termine *signe* verrà quindi conservato per indicare esclusivamente la combinazione dell'immagine acustica e dell'idea.

## 6. Conclusioni

Il modello *DIATERM*, illustrato nel presente contributo, è stato concepito con il fine ultimo di consentire una rappresentazione formale della fluttuazione terminologica che si verifica in alcune fasi della scienza, in quanto specchio di una concettualizzazione mobile e progressiva.

La necessità, tuttavia, di padroneggiare i meccanismi di reificazione e il linguaggio di formalizzazione OWL alla base del modello *DIATERM* rendono la costruzione di risorse diacroniche alquanto complessa per il terminologo digiuno di tali formalismi.

Per tale motivo è in corso di realizzazione una estensione di LexO, che consentirà allo studioso di creare risorse diacroniche in modo semplice e intuitivo. Tali risorse, compatibili con gli standard del web semantico e le tecnologie dei *Linked Data*, potranno essere messe a disposizione dell'intera comunità scientifica.

Un'ulteriore direzione di lavoro è rappresentata dall'integrazione in LexO dei prototipi di visualizzazione interattiva presentati nel contributo, i quali potranno aiutare lo studioso a comprendere a fondo le dinamiche spesso complesse che si instaurano tra teorizzazione e terminologia.

## Riferimenti bibliografici

- Batsakis, S. *et al.* 2017. "Temporal representation and reasoning in OWL 2". *Semantic Web* 8 (6): 981-1000.
- Bellandi, A., Giovannetti, E. e A. Weingart 2018. "Multilingual and Multiword Phenomena in a lemon Old Occitan Medico-Botanical Lexicon". *Information* 9 (3): 52.
- Bellandi, A. 2021. "LexO: an open-source system for managing OntoLex-Lemon resources". *Language Resources & Evaluation*. DOI: <https://doi.org/10.1007/s10579-021-09546-4>.

- Bezerianos, A. *et al.* 2010. "GeneaQuilts: A System for Exploring Large Genealogies". *IEEE Transactions on Visualization and Computer Graphics* 16 (6): 1073-81.
- Bourigault, D., Slodzian, M. 1999. "Pour une terminologie textuelle". *Terminologies nouvelles* 19: 29-32.
- Grewe, N. 2010. "A Generic Reification Strategy for n-ary Relations in DL". In *Proceedings of the 2<sup>nd</sup> Workshop of the GI-Fachgruppe "Ontologien in Biomedizin und Lebenswissenschaften" (OBML)*, eds. Herre H. *et al.* Mannheim, Germany, 2/2010. In IMISE-Report, Institute of Medical Informatics, Statistics and Epidemiology, University of Leipzig.
- Humbley, J. 2011. "Vers une méthode de terminologie rétrospective". *Langages* 183 (3): 51-62.
- Klein, M. CA., Fensel, D. 2001. "Ontology versioning on the Semantic Web". *Proceedings of the First International Conference on Semantic Web Working*, 75-91. California: CEUR-WS.org.
- Kuhn, T.S. 1970. *The structure of scientific revolutions*. Chicago: The University of Chicago Press (ed. orig. 1962).
- Manola, F., Miller, E. 2004. *RDF primer*. Technical report, W3C. <<http://www.w3.org/TR/2004/REC-rdf-primer-20040210/>> (2021-06-29).
- McCrae *et al.* 2017. "The OntoLex-Lemon Model: Development and Applications". *Proceedings of Electronic Lexicography in the 21<sup>st</sup> Century*: 19-21.
- Noy, N. *et al.* 2006. *Defining N-ary Relations on the Semantic Web*. W3C Working Group Note, aprile 2006.
- Piccini, S., Bellandi, A. e G. Benotto. 2016. "Formalizing and Querying a Diachronic Termino-Ontological Resource: the CLAVIUS Case Study". *Proceedings of From Digitization to Knowledge 2016 workshop (D2K)*, 38-41. Krakow: Linköping University Electronic Press, Linköpings universitet.
- Ruimy, N. *et al.* 2013. "Lessicografia Computazionale e Terminologia Saussuriana". In *Guida per un'edizione digitale dei manoscritti di Ferdinand de Saussure*, a cura di D. Gambarara, M. P. Marchese, 161-79. Alessandria: Edizioni dell'Orso.
- Slodzian, M. 2000. "L'émergence d'une terminologie textuelle et le retour du sens". In *Les sens en terminologie*, dir H. Béjoint e Ph. Thoiron, 61-85. Lyon: Presses Universitaires de Lyon.
- Van Campenhoudt, M. 1997. "Maille ou maillon: quand des terminographes négligent l'évolution de l'usage". In *Proceedings of the 5<sup>th</sup> scientific days Réseau: Lexicologie, Terminologie, Traduction* (Agence Universitaire de la Francophonie), *La mémoire des mots*, eds. A. Clas *et al.*, 251-272. Tunis, 25-27 septembre: 251-272.
- Welty, Ch., Fikes, R., Makarios, S. 2006. "A reusable ontology for fluents in owl". *FOIS* 150: 226-36.
- Wilkinson *et al.* 2016. "The FAIR Guiding Principles for scientific data management and stewardship". *Scientific data* 3. doi: <https://doi.org/10.1038/sdata.2016.18>
- Wüster, E. 1968. *The Machine Tool. An Interlingual dictionary of basic concepts*. London: Technical Press.
- Zanola, M. T. 2014. *Arts et métiers au XVIII<sup>e</sup> siècle. Études de terminologie diachronique*. Paris: L'Harmattan.