The United Nations Document Ontology

Deliverable, March 19, 2017

# Abstract

In this document we introduce the United Nations Document Ontology (UNDO), i.e. an OWL 2 DL ontology that aims at providing a framework for the description of all the entities mentioned in United Nations documents, and the relations that can exists among them.

# Introduction

The parliamentary, normative and legal documents published by the United Nations system of organisations, from now on “UN documents”, are full of references to real-world objects and concepts, such as other documents, people, organisations, legal terms, and roles. Such entities are linked with the actual content of the document, even by means of different terms being the UN documents often issued in several official languages.

The United Nations Document Ontology (UNDO, <https://w3id.org/akn/ontology/undo>; GitHub repository available at <https://github.com/essepuntato/undo>) is an OWL 2 DL ontology that aims at providing a framework for the description of all these entities and the relations that can exists among them. The idea behind the development of this model is to provide a common framework to be used and, eventually, extended by the various agencies of the United Nations for sharing data about documents and their content in RDF format in an interchangeable way.

The rest of the document is organised as follows. In Methods and Material we introduce all the methods and material that have been used for the development of the ontology. In The United Nations Document Ontology (UNDO) we summarise the ontology, highlighting its most important classes and some design choices. In Exemplar of usage we introduce some example of how the ontology can be used for describing real-case scenarios. In Querying the ontology we show how to convert natural language questions into formal query so as to retrieve data according to UNDO. Finally, in Conclusions and future directions, we conclude the document introducing some future works.

# Methods and Material

In this section we illustrate all the methods and material that have supported the development of UNDO.

## Simplified Agile Methodology for Ontology Development

The Simplified Agile Methodology for Ontology Development (SAMOD, <https://github.com/essepuntato/samod>) is a novel agile methodology for the development of ontologies, partially inspired to the Test-Driven Development process in Software Engineering and to existing agile ontology development methodologies such as eXtreme Design (XD). SAMOD is organised in three simple steps within an iterative process that focuses on creating well-developed and documented models by using exemplars of data, so as to produce ontologies that are always ready-to-be-used and easily-understandable by humans (i.e. the possible customers) without spending a lot of effort. It has been used to develop UNDO since the beginning.

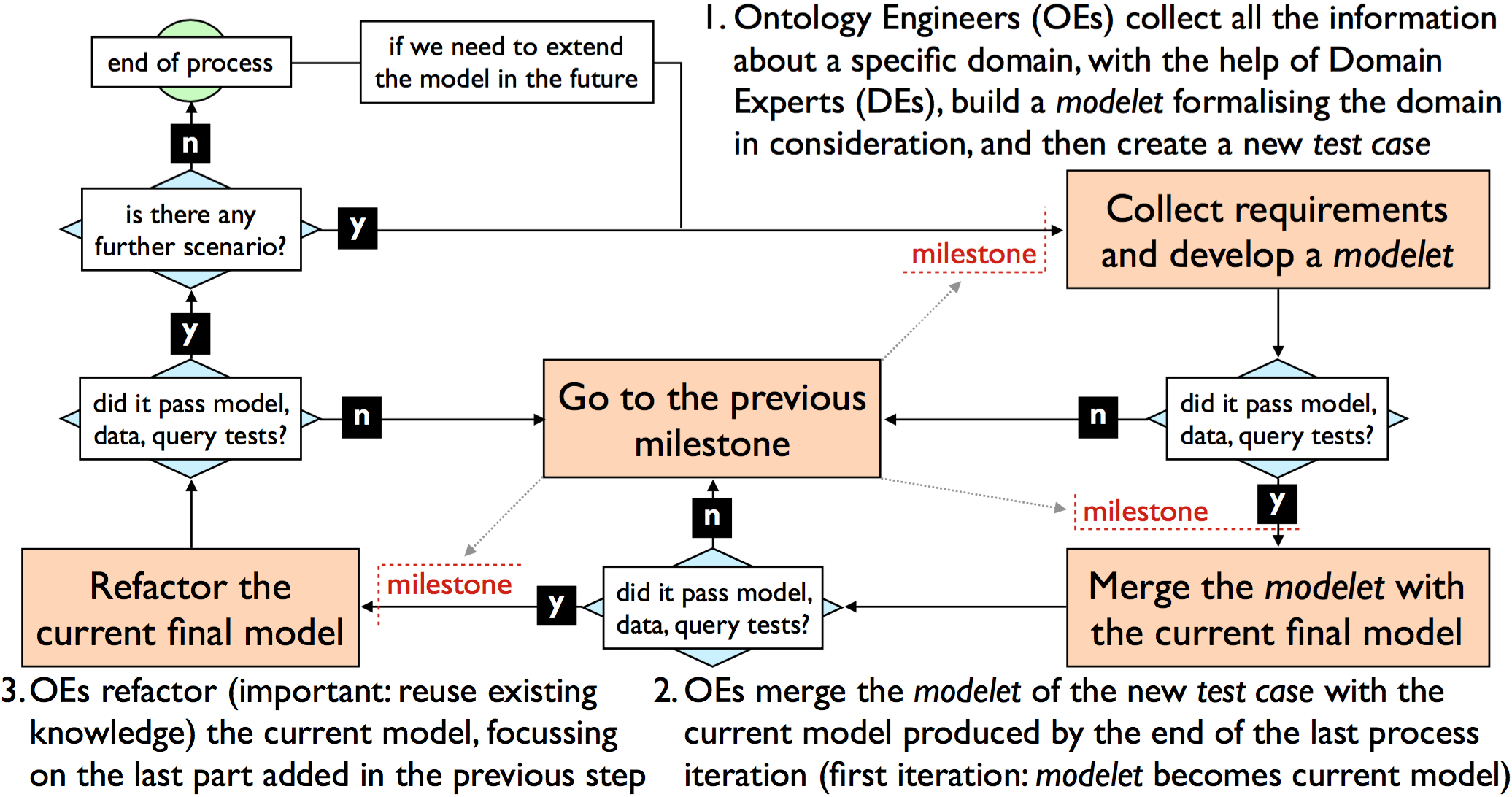


Figure 1 A summary of the three steps of SAMOD.

Each step of the methodology, summarised in Figure 1, ends with the release of a snapshot of the current state of the process called milestone, and involves one or more ontology engineers, the experts in semantic technologies and ontology development tools and languages, or OEs, and some domain experts, the experts of the particular domain to be modelled, in the following way:

1. the ontology engineers collected all the information about a specific domain, with the help of domain experts. Then, ontology engineers build a small monolithic model, called *modelet*, formalising the domain in consideration, following specific ontology development principles. Finally, a new *test case* is created – i.e., a set of resources that includes the modelet, some *exemplar data* and *query* to be answered, that must be appropriately tested by means of formal tools, such as reasoners. If everything works fine, a milestone is released and the process continues, otherwise the ontology engineers have to go back to the previous milestone;
2. the ontology engineer merges the modelet of the new test case with the *current model* produced by the end of the last iteration of the process, and consequently update and check all the test cases developed in the past so as to include the new current model. If everything works fine, a milestone is released and the process continues, otherwise the ontology engineers have to go back to the previous milestone;
3. the ontology engineer refactors the current model, in particular focussing on the last part added in the previous step, taking into account *good practices* for ontology development processes. If everything works fine, a milestone is released, otherwise the ontology engineers have to go back to the previous released milestone. Thus, in case there is another motivating scenario to be addressed, then the process is iterated, otherwise it stops.

## Live OWL Documentation Environment

The Live OWL Documentation Environment (LODE, <http://www.essepuntato.it/lode>) is a service that automatically extracts classes, object properties, data properties, named individuals, annotation properties, general axioms and namespace declarations from an OWL and OWL2 ontology, and renders them as ordered lists, together with their textual definitions, in a human-readable HTML page designed for browsing and navigation by means of embedded links. This LODE service is an open source development, and can be freely used. It may be used in conjunction with content negotiation to display this human-readable version of an OWL ontology when the user accesses the ontology using a web browser, or alternatively to deliver the OWL ontology itself when the user accesses the ontology using an ontology editing tool such as Protégé and NeOn Toolkit.

In the context of the development of UNDO, Graffoo has been used to produce the HTML documentation of the ontology by extrapolating their annotations.

## Graphical Framework For OWL Ontologies

The Graphical Framework for OWL Ontologies (Graffoo, <http://www.essepuntato.it/graffoo>), is an open source tool that can be used to present the classes, properties and restrictions within OWL ontologies, or sub-sections of them, as clear and easy-to-understand diagrams. The advantages of using such a Graffoo diagram are, thus, that it displays the logical relationships between elements of an ontology, or a sub-section of an ontology, in a manner that is relatively straightforward to understand, once one has grasped the meaning of the different elements of a Graffoo diagram. These elements are fully presented in the official specification (<http://www.essepuntato.it/graffoo/specification/current.html>).

In the context of the development of UNDO, it has been used to create the various modelets (see Simplified Agile Methodology for Ontology Development) and all the diagrams of the ontology.

## Diagrams Transformation into OWL

The Diagrams Transformation inTo OWL (DiTTO, <http://www.essepuntato.it/ditto>) is a Web application that is able to translate diagrams expressed in either E/R crow’s foot notation or Graffoo and created with yEd (<http://www.yworks.com/en/products_yed_about.html>) into OWL ontologies. If one chooses to specify E/R diagrams, DiTTO allows one to choose what E/R semantics to apply for the transformation according to three alternative conversion strategies, which depends on the application of two assumptions:

* global semantics (GS) is a characteristic of OWL ontologies (but not typically of E/R), and has the effect of unifying the formal interpretation of domain and range axioms, property characteristics, and all the restrictions that act at a global level. When GS does not hold, one is not allowed to assume such unification, even when the axioms regard two constants with the same name;
* unique name assumption (UNA), which is a characteristic in E/R semantics (but not of OWL), and whose consequence is that two objects named differently always refer to different entities in the world.

In the context of the development of UNDO, DiTTO has been used to convert the Graffoo diagrams of the modelets into OWL automatically.

## A Light Legal Ontology on TLCs

Akoma Ntoso (<https://w3id.org/akn>) is an open XML standard for parliamentary, legislative and judiciary documents. The Akoma Ntoso specifications, instead of forcing document authors to stick to one particular ontology, specify a very broad set of general guidelines that an ontology should conform to in order to be used in conjunction with Akoma Ntoso documents. These informal guidelines are mentioned with the name of non-ontology, and are implicitly referred by with specific XML tags (prefixed by “TLC” or “FRBR”) in Akoma Ntoso documents, each referring to a particular (an informal) top level class (TLC).

ALLOT (A Light Legal Ontology On TLCs, <https://w3id.org/akn/ontology/allot>) implements the Akoma Ntoso TLCs as a formal OWL 2 DL ontology, so as to model the entities that are present in legal documents (i.e., people, events, roles, etc.) with the main purpose of linking together datasets based, potentially, on different ontologies. ALLOT formalises the Akoma Ntoso TLCs in a set of OWL classes, and has developed following the Akoma Ntoso non-ontology, its associated guidelines and best practices.

All the material related with the ontology is available on the ALLOT GitHub repository at <https://github.com/essepuntato/allot>. The current version of ALLOT is actually based on a previous implementation named ALLOT Core and available at <http://akn.web.cs.unibo.it/allot/core.owl>. The ALLOT ontology is meant to be used to describe in detail the documental and non-documental references (i.e. any entity) present in Akoma Ntoso documents. ALLOT can also be used to bridge knowledge bases extracted from Akoma Ntoso documents to knowledge bases that use other ontologies such as Metalex (<http://justinian.leibnizcenter.org/MetaLex/metalex-cen.owl>).

In the context of the development of UNDO, ALLOT has been used as basic ontology to extend so as to be compliant with Akoma Ntoso document specification.

## Other related ontologies

For describing a domain concerning documents, and the entities they describe, other relevant ontologies are relevant and deserve to be properly introduced. In the following subsections, we provide a quick introduction of those that have been directly reused in UNDO.

### TVC

Time-indexed Value in Context (TVC, <http://www.essepuntato.it/2012/04/tvc>) is an ontology pattern that allows one to describe scenarios in which someone (e.g., a person) has a value (e.g., a particular role) during a particular time and for a particular context.

### Time Interval

Time Interval (TI, <http://www.ontologydesignpatterns.org/cp/owl/timeinterval.owl>) is an ontology pattern that enables the description of period of times characterised by a starting date and an ending date.

### Web Annotation Ontology

Web Annotation Ontology (<https://www.w3.org/ns/oa>) is a set of RDF classes, predicates and named entities that are used by the Web Annotation Data Model for creating annotations in RDF.

### DCTerms

Dublin Core Metadata Terms (DCTerms, <http://purl.org/dc/terms/>) is an ontology implementing all the metadata terms maintained by the Dublin Core Metadata Initiative, including properties, vocabulary encoding schemes, syntax encoding schemes, and classes.

### FaBiO

The FRBR-align Bibliographic Ontology (FaBiO, <http://purl.org/spar/fabio>) is an ontology for recording and publishing on the Semantic Web bibliographic records of scholarly endeavours.

### FOAF

Friend Of A Friend (FOAF, <http://xmlns.com/foaf/0.1/>) is an ontology for describing people and their relations with other people, documents, and other information objects.

### FRBR DL

Functional Requirement for Bibliographic Records DL (FRBR DL, <http://purl.org/spar/frbr>) is an expression in OWL 2 DL of the basic concepts and relations described in the IFLA report on the Functional Requirements for Bibliographic Records (FRBR), also described in Ian Davis's RDF vocabulary.

### ISO 639-1

ISO 639-1 (<http://id.loc.gov.vocabulary/iso693-1>) is a vocabulary describing the first part of the ISO 639 international-standard language-code family.

### LKIF Core

Two different ontologies defined within the framework LKIF Core has been considered:

* LKIF Core: Action (<http://www.estrellaproject.org/lkif-core/action.owl>) is an ontology for representing actions in general, i.e. processes which are performed by some agent;
* LKIF Core: Role (<http://www.estrellaproject.org/lkif-core/role.owl>) is an ontology for describing typology of roles (epistemic roles, functions, person roles, organisation roles).

### PRO

The Publishing Roles Ontology (PRO, <http://purl.org/spar/pro>) is an ontology describing possible roles in the publication process, or in other scholarly activities or situations, held by particular agent;

### PSO

The Publishing Status Ontology (PSO, <http://purl.org/spar/pso>) is an ontology for characterizing the publication status of a document or other publication entity at each of the various stages in the publishing process;

### PWO

The Publishing Workflow Ontology (PWO, <http://purl.org/spar/pwo>) is a simple ontology written in OWL 2 DL for the characterization of the main stages in the workflow associated with the publication of a document (e.g. being written, under review, XML capture, page design, publication to the Web).

### SKOS

The Simple Knowledge Organization System (SKOS, <https://www.w3.org/TR/skos-reference/>) is a common data model for sharing and linking knowledge organization systems via the Web.

# The United Nations Document Ontology (UNDO)

In this section we introduce the main development choices used for the United Nations Document Ontology (UNDO), i.e. an OWL 2 DL ontology that aims at providing a framework for the description of all the entities mentioned in United Nations documents, and the relations that can exists among them.

Currently, UNDO is able to describe the following entities:

* documents and their versions (e.g. in different languages);
* several relevant document types, organised in a taxonomy;
* relations, with the possibility of defining their semantics dynamically (outside the scope of the TBox of the ontology) as individual of the class defining concepts;
* annotations, in order to link any RDF description to a particular document;
* terms, i.e. tokens as they can appear in a text;
* concepts, i.e. any notion or idea that can be used to provide a meaning to terms and a specific semantics to relations;
* values that can change in time and context (such as agents’ roles and document statuses);
* workflows and their executions.

While the ontology has been developed from scratch, we have also reused existing and well-known models so as to make UNDO interoperable in different context. So as to regulate the way the entities defined in such external models have been reused, we have applied the following guidelines:

* some of these ontologies (i.e. ALLOT, PWO, TVC, Time Interval, and Web Annotation Ontology) have been imported as a whole (by means of the property owl:imports), since some of them have entities that are directly reused in UNDO for providing a description of the domain in consideration (e.g. we use allot:hasRealization to link a document to its version in a specific language);
* some ontological entities defined in external models (i.e. DCTerms, FOAF, and ISO 639-1, referred via rdfs:isDefinedBy) have been reused in UNDO (e.g. dcterms:language) without importing the original models since they have not been defined formally as OWL 2 DL ontologies;
* other ontological entities defined in external models (i.e. FaBiO, FRBR DL, LKIF Core, PRO, PSO, and SKOS, referred via rdfs:isDefinedBy), which are proper OWL 2 DL ontologies, have been included for the sake of aligning UNDO with other relevant and existing models.

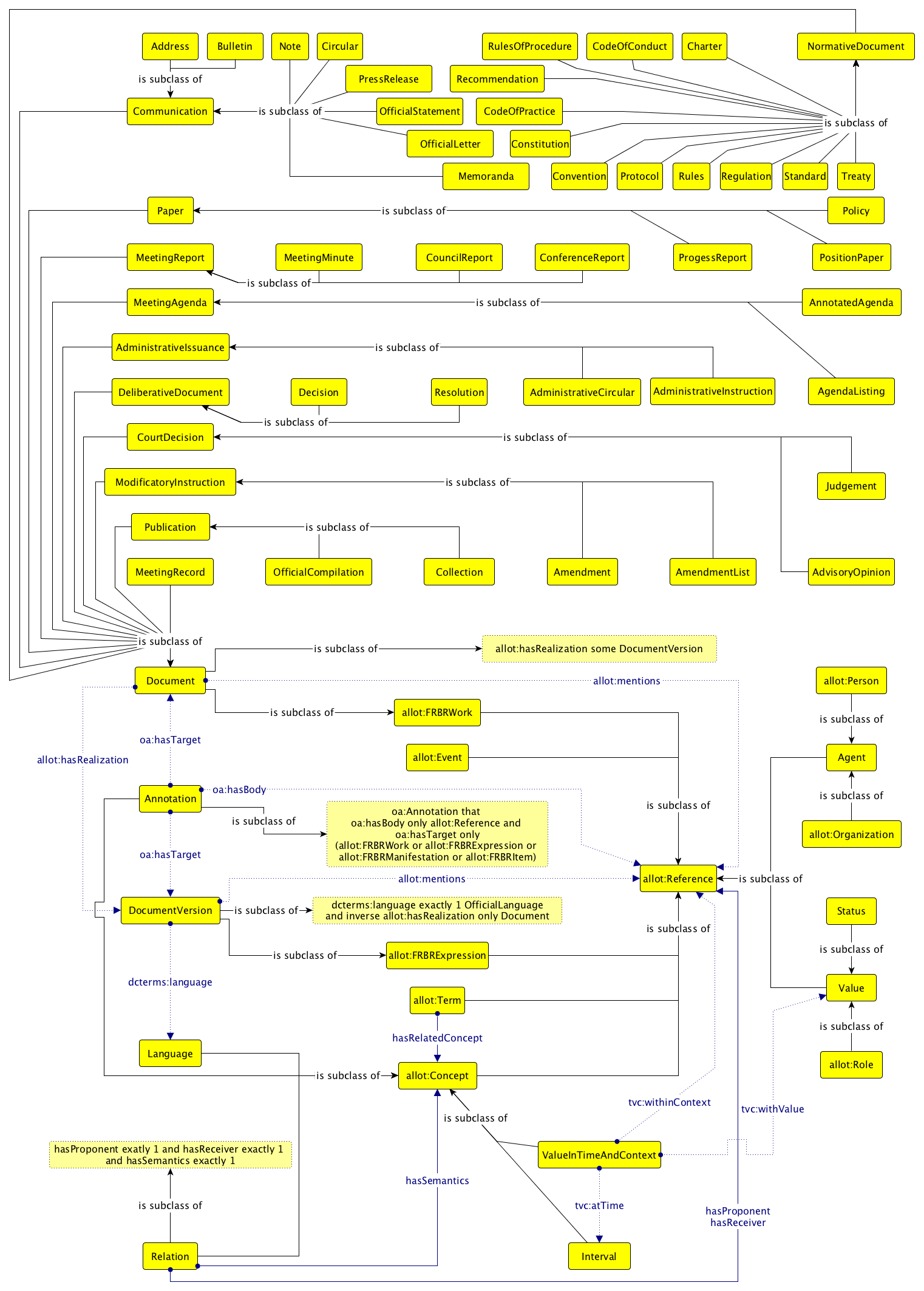


Figure 2 The Graffoo diagram of the United Nations Document Ontology (UNDO).

The diagram in Figure 2 introduces the version of UNDO dated Match 8, 2017, while the new version will be released soon, so as to include also the full specification for describing workflows – which will be largely discussed in the new version of this deliverable. In the next subsection we provide a quick overview of all the entities defined.

## Documents and their versions

The parliamentary, normative and legal documents published by the United Nations system of organisations are full of references to real-world objects and concepts, such as other documents, people, organisations, legal terms, and roles. Such entities are linked with the actual content of the document, even by means of different terms being the UN documents often issued in several official languages.

These entities and relations are, thus, implicitly defined by the content of the document in consideration, and typically can concern at least one of the following layers:

* document-oriented layer, which allows the creation of direct links (via allot:mentions) from a document (undo:Document) or a specific version of it (undo:DocumentVersion) to other relevant entities that are mentioned within;
* entity-oriented layer, which enables the identification of entities cited in the document content (e.g. organisations) acting with other relevant entities (people, events, organizations, etc.), described by the class undo:Relation;
* content-oriented layer, which provides a specific semantics (via undo:hasRelatedConcept) to all the terms (allot:Term) identified in the document content.

In addition, each document (or any of its parts) can be explicitly annotated (using instances of undo:Annotation) by someone with all these relations in order to create a formal connection between the text and the entities its introduces.

## Time and context

Several kinds of objects can be involved in situations (undo:ValueInTimeAndContext) describing them as holding a certain value (tvc:withValue) associated for a specific interval (tvc:atTime) or according to a specific context (tvc:withinContext). For instance, agents having a particular role (e.g. being the president of the Italy from 2006 to 2015) or documents holding particular status (e.g. a document that has been under-review from September 2016 to October 2016) in a specific time and/or related to a particular context are examples for these situations.

These kinds of scenarios involve at least four specific kinds of entities:

\* the entity holding such value (e.g. an agent or a document);

\* the value held (e.g. a role or a status);

\* the time defining when such value is held (e.g. from 2006 to 2015);

\* the context to which the scenario applies (Italy).

## Workflows and their executions

Keeping track of the processes concerning the creation and modification of documents is a crucial task to address in the legal and legislative domain. Each of these processes, commonly called workflow, is actually composed by a sequence of steps. Each step is responsible to produce some outputs (e.g. a review) starting from some inputs (e.g. a document).

Workflows can be described from two different points of view. On the one hand, there is the declaration of the workflow schema, or simply the workflow (pwo:Workflow), which is how a specific process (e.g. the publication of a document) is actually organised in sequential steps (pwo:Step), each having a specific type (e.g. write, review, revise), involving different kinds of actors, and producing, in the end, a particular output. On the other hand, each particular execution of a workflow (pwo:WorkflowExecution) is a specific entity per se, it is usually composed by sets of actions (pwo:Action), and each set of actions is addressing a particular step defined in the workflow in consideration.

# Exemplar of usage

UNDO can be used for modelling several scenarios that are typical to the documents published by the United Nations (e.g. UN Resolutions). In the following subsections we introduce some of them, and we accompany them with exemplar instantiations.

All the examples are provided in Turtle format [XX], which is a particular linearization for RDF statements. The prefixes that are used in all the examples introduced in the following sub-sections are defined as follows:

@prefix : <https://w3id.org/akn/data/undo/> .

@prefix undo: <https://w3id.org/akn/ontology/undo/> .

@prefix allot: <https://w3id.org/akn/ontology/allot/> .

@prefix owl: <http://www.w3.org/2002/07/owl#> .

@prefix dcterms: <http://purl.org/dc/terms/> .

@prefix lan: <http://id.loc.gov/vocabulary/iso639-1/> .

@prefix oa: <http://www.w3.org/ns/oa#> .

@prefix ti: <http://www.ontologydesignpatterns.org/cp/owl/timeinterval.owl#> .

@prefix tvc: <http://www.essepuntato.it/2012/04/tvc/> .

## Scenario 1

**Description:** In the preface of the UN resolution A/RES/50/100 of 2 February 1996 there are references to:

* the UN resolution A/RES/47/180 of 22 December 1992;
* the Government of Turkey;
* the United Nations Conference on Human Settlements (Habitat II).

**Turtle:**

:a-res-50-100-1996-02-02 a undo:Resolution ;

allot:mentions

:a-res-47-180-1992-12-22 ,

:government-of-turkey ,

:habitat-ii .

:a-res-47-180-1992-12-22 a undo:Resolution .

:government-of-turkey a allot:Organization .

:habitat-ii a allot:Event .

## Scenario 2

**Description:** According to the content of the UN resolution A/RES/50/100 of 2 February 1996, the General Assembly:

* recalls the resolution A/RES/47/180 of 22 December 1992;
* reiterates its gratitude to the Government of Turkey;
* notes the progress that has been made so far in the preparations for the United Nations Conference on Human Settlements (Habitat II).

**Turtle:**

:a-res-50-100-1996-02-02 a undo:Resolution .

:a-res-47-180-1992-12-22 a undo:Resolution .

:general-assembly a allot:Organization .

:organization a allot:Organization .

:government-of-turkey a allot:Organization .

:habitat-ii a allot:Event .

:recall a allot:Concept .

:reiterate-its-gratitude a allot:Concept .

:note-the-progress a allot:Concept .

:relation-1 a undo:Relation ;

undo:hasProponent :general-assembly ;

undo:hasSemantics :recall ;

undo:hasReceiver :a-res-47-180-1992-12-22 .

:relation-2 a undo:Relation ;

undo:hasProponent :general-assembly ;

undo:hasSemantics :reiterate-its-gratitude ;

undo:hasReceiver :government-of-turkey .

:relation-3 a undo:Relation ;

undo:hasProponent :general-assembly ;

undo:hasSemantics :note-the-progress ;

undo:hasReceiver :habitat-ii .

:annotation-1-a-res-50-100-1996-02-02 a undo:Annotation ;

oa:hasBody :relation-1 ;

oa:hasTarget :a-res-50-100-1996-02-02 .

:annotation-2-a-res-50-100-1996-02-02 a undo:Annotation ;

oa:hasBody :relation-2 ;

oa:hasTarget :a-res-50-100-1996-02-02 .

:annotation-3-a-res-50-100-1996-02-02 a undo:Annotation ;

oa:hasBody :relation-3 ;

oa:hasTarget :a-res-50-100-1996-02-02 .

## Scenario 3

**Description:** The UN resolution A/RES/50/100 has been published in six different languages. The term "decides" has been used twice in the English version, and, in parallel, the term "decide" has been used twice in the Spanish version. Both the terms actually refers to the same higher-level concept of taking a decision, even if this aspect has not been made formally explicit in the document itself.

**Turtle:**

:a-res-50-100-1996-02-02 a undo:Resolution ;

allot:hasRealization

:a-res-50-100-1996-02-02-en ,

:a-res-50-100-1996-02-02-es .

:a-res-50-100-1996-02-02-en a undo:DocumentVersion ;

dcterms:language lan:en .

:a-res-50-100-1996-02-02-es a undo:DocumentVersion ;

dcterms:language lan:es .

:decides a allot:Term ;

undo:hasRelatedConcept :taking-a-decision .

:annotation-1-a-res-50-100-1996-02-02-en a undo:Annotation ;

oa:hasBody :decides ;

oa:hasTarget :a-res-50-100-1996-02-02-en .

:annotation-2-a-res-50-100-1996-02-02-en a undo:Annotation ;

oa:hasBody :decides ;

oa:hasTarget :a-res-50-100-1996-02-02-en .

:decide a allot:Term ;

undo:hasRelatedConcept :taking-a-decision .

:annotation-1-a-res-50-100-1996-02-02-es a undo:Annotation ;

oa:hasBody :decide ;

oa:hasTarget :a-res-50-100-1996-02-02-es .

:annotation-2-a-res-50-100-1996-02-02-es a undo:Annotation ;

oa:hasBody :decide ;

oa:hasTarget :a-res-50-100-1996-02-02-es .

:taking-a-decision a allot:Concept .

## Scenario 4

**Description:** Giorgio Napolitano has been President of Italy for the first time from the 15th of May 2006 to the 22th of April 2013, when he resigned. He has been elected again on the 22th of April 2013, and he has been President of Italy for the second time by the 14th of January 2015, when he resigned for the second time. He has been substituted by Sergio Mattarella, who starts his mandate as President of Italy on the 3rd of February 2015, formerly judge of the Constitutional Court of Italy from the 11th of October 2011 to the 2nd of February 2015.

**Turtle:**

:giorgio-napolitano a undo:Agent ;

tvc:hasValue

:giorgio-napolitano-president-of-italy-2006-2013 ,

:giorgio-napolitano-president-of-italy-2013-2015 .

:giorgio-napolitano-president-of-italy-2006-2013 a undo:ValueInTimeAndContext ;

tvc:withValue :president ;

tvc:atTime :2006-2013 ;

tvc:withinContext :italy .

:giorgio-napolitano-president-of-italy-2013-2015 a undo:ValueInTimeAndContext ;

tvc:withValue :president ;

tvc:atTime :2013-2015 ;

tvc:withinContext :italy .

:president a allot:Role .

:2006-2013 a undo:Interval ;

ti:hasIntervalStartDate "2006-05-15T00:00:00"^^xsd:dateTime ;

ti:hasIntervalEndDate "2013-04-22T00:00:00"^^xsd:dateTime .

:2013-2015 a undo:Interval ;

ti:hasIntervalStartDate "2013-04-22T00:00:00"^^xsd:dateTime ;

ti:hasIntervalEndDate "2015-01-14T00:00:00"^^xsd:dateTime .

:italy a allot:Reference .

:sergio-mattarella a undo:Agent ;

tvc:hasValue

:sergio-mattarella-judge-constitutional-court-of-italy-2011-2015 ,

:sergio-mattarella-president-of-italy-since-2015.

:sergio-mattarella-judge-constitutional-court-of-italy-2011-2015 a undo:ValueInTimeAndContext ;

tvc:withValue :judge ;

tvc:atTime :2011-2015 ;

tvc:withinContext :constitutional-court-of-italy .

:sergio-mattarella-president-of-italy-since-2015 a undo:ValueInTimeAndContext ;

tvc:withValue :president ;

tvc:atTime :since-2015 ;

tvc:withinContext :italy .

:2011-2015 a undo:Interval ;

ti:hasIntervalStartDate "2011-10-11T00:00:00"^^xsd:dateTime ;

ti:hasIntervalEndDate "2015-02-03T00:00:00"^^xsd:dateTime .

:since-2015 a undo:Interval ;

ti:hasIntervalStartDate "2015-02-03T00:00:00"^^xsd:dateTime .

:constitutional-court-of-italy a allot:Reference .

:judge a allot:Role .

## Scenario 5

**Description:** A document has been in drafting from the 14th of May to the 14th of June 2016. Then, it has been under revision for the following two months, and finally published on the 1st of September 2016.

**Turtle:**

:a-document a undo:Document ;

tvc:hasValue

:document-in-drafting ,

:document-under-revision ,

:document-published .

:document-in-drafting a undo:ValueInTimeAndContext ;

tvc:withValue :in-drafting ;

tvc:atTime :in-drafting-interval .

:document-under-revision a undo:ValueInTimeAndContext ;

tvc:withValue :under-revision ;

tvc:atTime :under-revision-interval .

:document-published a undo:ValueInTimeAndContext ;

tvc:withValue :published ;

tvc:atTime :published-interval .

:in-drafting-interval a undo:Interval ;

ti:hasIntervalStartDate "2016-05-14T00:00:00"^^xsd:dateTime ;

ti:hasIntervalEndDate "2016-06-14T00:00:00"^^xsd:dateTime .

:under-revision-interval a undo:Interval ;

ti:hasIntervalStartDate "2016-06-14T00:00:00"^^xsd:dateTime ;

ti:hasIntervalEndDate "2016-08-14T00:00:00"^^xsd:dateTime .

:published-interval a undo:Interval ;

ti:hasIntervalStartDate "2016-09-01T00:00:00"^^xsd:dateTime .

:in-drafting a undo:Status .

:under-revision a undo:Status .

:published a undo:Status .

# Querying the ontology

UNDO can be used for answering several questions related to the content of the documents published by the United Nations (e.g. UN Resolutions). In the following subsections we introduce some of them, and we accompany them with exemplar SPARQL queries.

The prefixes that are used in all the SPARQL queries provided below are defined as follows:

PREFIX : <https://w3id.org/akn/data/undo/>

PREFIX allot: <https://w3id.org/akn/ontology/allot/>

PREFIX oa: <http://www.w3.org/ns/oa#>

PREFIX owl: <http://www.w3.org/2002/07/owl#>

PREFIX undo: <https://w3id.org/akn/ontology/undo/>

## Question 1

**Natural language question:** What are all the documents that refer to the UN resolution A/RES/47/180 of 22 December 1992?

**SPARQL:**

SELECT DISTINCT ?document

WHERE {

?document allot:mentions :a-res-47-180-1992-12-22 .

}

## Question 2

**Natural language question:** What are all the entities referred by the UN resolution A/RES/50/100 of 2 February 1996?

**SPARQL:**

SELECT DISTINCT ?entity ?type

WHERE {

:a-res-50-100-1996-02-02 allot:mentions ?entity .

?entity a ?type .

FILTER (?type != owl:NamedIndividual)

}

## Question 3

**Natural language question:** What are the documents that describes the General Assembly recalling the resolution A/RES/47/180 of 22 December 1992?

**SPARQL:**

SELECT DISTINCT ?document

WHERE {

?rel a undo:Relation ;

undo:hasProponent :general-assembly ;

undo:hasSemantics :recall ;

undo:hasReceiver :a-res-47-180-1992-12-22 .

?ann a undo:Annotation ;

oa:hasBody ?rel ;

oa:hasTarget ?document

}

## Question 4

**Natural language question:** What are all the terms that actually refer to the concept of taking a decision, and in which documents have been used?

**SPARQL:**

SELECT DISTINCT ?term ?document

WHERE {

?term a allot:Term ;

undo:hasRelatedConcept :taking-a-decision .

?ann a undo:Annotation ;

oa:hasBody ?term ;

oa:hasTarget ?document

}

## Question 5

**Natural language question:** What are the current roles of all the agents?

**SPARQL:**

SELECT DISTINCT ?agent ?role ?context

WHERE {

?agent a undo:Agent ;

tvc:hasValue ?value .

?value a undo:ValueInTimeAndContext ;

tvc:withValue ?role ;

tvc:withinContext ?context .

FILTER NOT EXISTS {

?value tvc:atTime/ti:hasIntervalEndDate ?time

}

?role a allot:Role .

}

## Question 6

**Natural language question:** What roles have been held by Sergio Mattarella during his life?

**SPARQL:**

SELECT DISTINCT ?role ?context ?start ?end

WHERE {

:sergio-mattarella tvc:hasValue ?value .

?value a undo:ValueInTimeAndContext ;

tvc:withValue ?role ;

tvc:withinContext ?context ;

tvc:atTime ?time .

?time a undo:Interval ;

ti:hasIntervalStartDate ?start .

OPTIONAL {

?time ti:hasIntervalEndDate ?end

}

?role a allot:Role .

}

## Question 7

**Natural language question:** How many times an agent has been assigned to a specific role?

**SPARQL:**

SELECT DISTINCT ?agent ?role ?context (count(?agent) AS ?count)

WHERE {

?agent a undo:Agent ;

tvc:hasValue ?value .

?value a undo:ValueInTimeAndContext ;

tvc:withValue ?role ;

tvc:withinContext ?context .

?role a allot:Role .

}

GROUP BY ?agent ?role ?context

## Question 8

**Natural language question:** Which roles are involved in particular contexts?

PREFIX allot: <https://w3id.org/akn/ontology/allot/>

PREFIX undo: <https://w3id.org/akn/ontology/undo/>

PREFIX tvc: <http://www.essepuntato.it/2012/04/tvc/>

SELECT DISTINCT ?context ?role

WHERE {

?value a undo:ValueInTimeAndContext ;

tvc:withValue ?role ;

tvc:withinContext ?context .

?role a allot:Role .

}

## Question 9

**Natural language question:** What are all the statuses that have been associated to all the documents?

**SPARQL:**

PREFIX allot: <https://w3id.org/akn/ontology/allot/>

PREFIX undo: <https://w3id.org/akn/ontology/undo/>

PREFIX ti: <http://www.ontologydesignpatterns.org/cp/owl/timeinterval.owl#>

PREFIX tvc: <http://www.essepuntato.it/2012/04/tvc/>

SELECT DISTINCT ?document ?status ?start ?end

WHERE {

?document a undo:Document ;

tvc:hasValue ?value .

?value a undo:ValueInTimeAndContext ;

tvc:withValue ?status ;

tvc:atTime ?time .

?time a undo:Interval ;

ti:hasIntervalStartDate ?start .

OPTIONAL {

?time ti:hasIntervalEndDate ?end

}

?status a undo:Status .

}

# Conclusions and future directions

In this document we have provided an overview of the United Nations Document Ontology (UNDO), i.e. an OWL 2 DL ontology that aims at providing a framework for the description of all the entities mentioned in United Nations documents, and the relations that can exists among them. The idea behind the development of this model is to provide a common framework to be used and, eventually, extended by the various agencies of the United Nations for sharing data about documents and their content in RDF format in an interchangeable way.

There are several works that can be done in the future in the context of UNDO and related entities, so as to guarantee its broad usage within the United Nations:

* aligning UNDO to foundational ontologies, such as the Basic Formal Ontology and DOLCE;
* developing a URI resolver for Akoma Ntoso TLC URIs so as to associate the HTTP URI of the related resource available in an RDF-based knowledge base of the United Nations (compliant with UNDO);
* studying and implementing an algorithm (e.g. based on XSLT) that takes Akoma Ntoso documents as input and returns a set of RDF statements compliant with UNDO by interpreting Akoma Ntoso URIs and metadata sections;
* extending the United Nations Document Ontology so as to handle other aspects concerning the document content, e.g. organisations, languages, actions, document structures and rhetorics, and argumentation;
* developing taxonomies (e.g. defining the meaning of all the possible terms related to actions) that can be reused in the context of UNDO.