A geologic timescale ontology and service

Article *in* Earth Science Informatics · March 2014

DOI: 10.1007/s12145-014-0170-6

2 authors:

Simon Jonathan David Cox
The Commonwealth Scientific and Industrial Research Organisation

Stephen Miller Richard
US Geoscience Information Network

[revised 2014-05-14]

A geologic timescale ontology and service

S J D Cox[1] & S M Richard[2]

[1]CSIRO Land and Water, Highett, Vic. Australia simon.cox@csiro.au
[2]Arizona Geologic Survey, Tucson, Az. USA steve.richard@azgs.az.gov

Abstract

We have developed an OWL ontology for the geologic timescale, derived from a Unified Modeling Language (UML) model that formalized the practice of the International Commission for Stratigraphy (ICS) (Cox and Richard 2005). The UML model followed the ISO/TC 211 modeling conventions, and was the basis for an XML implementation that was integrated into GeoSciML 3.0. The OWL ontology is derived using rules for generating OWL ontologies from ISO-conformant UML models, as provided in a (draft) standard from ISO/TC 211. The basic ontology is also aligned with SKOS to allow multilingual labels, and to enable delivery through a standard vocabulary interface. All versions of the International Stratigraphic Chart from 2004 to 2013 have been encoded using the ontology. Following ICS practice, the elements of the timescale retain the same identifiers across the multiple versions, though the information describing each geochronologic unit evolves with the versions of the timescale. The timescales are published through multiple web interfaces and APIs.

Keywords:
OWL, SKOS, UML, stratigraphy, timescale, chronostratigraphy, reference system, versioning

1. Introduction

The geologic timescale is a fundamental reference system required for geology. Correlation of geologic features between widely separated locations is accomplished with the assistance of this timescale, so access to a precise description of the timescale is essential. Contemporary semantic web technologies can be used to enable this.

The conventional timescale is a hierarchical ordinal reference system, characterized by a nested system of ordered intervals. Principles for its formalization are provided by the International Commission for Stratigraphy (ICS)[1] (Gradstein et al. 2005).

The time scale for Phanerozoic time is based on defining a sequence of geologic events of sufficient magnitude that they have left material evidence in the rock record that can be recognized globally. The named intervals of the geologic time scale are bounded by these events. Because the time scale originated with the study of sequences of fossils in sequences of stacked sedimentary rocks, most of the boundary events are associated with changes in fossil assemblages interpreted to reflect evolutionary developments with global impact. The primary focus of the ICS is on establishing reference localities, each known as a Global Stratotype Section and Point (GSSP), where material corresponding to interval boundaries is available. These may then be used to calibrate the ages of these boundaries and thus define the timescale quantitatively. The definition of the geologic timescale in the Phanerozoic depends, therefore, on a set of precise relationships between certain abstract concepts (time intervals and boundaries), their manifestation in geologic sequences, and observations on physical occurrences of these.

For the Precambrian, which is the long period of Earth history preceding the waxing of abundant fossil taxa in the sedimentary record, clearly no biostratigraphic subdivisions can be defined. Large scale patterns in the geochemical development of the Earth system may be used to broadly define rock assemblages of older and younger Precambrian time, but it was not until the advent of isotope-based geochronology that the sequential relationships of Precambrian rocks could be worked out on a global scale. As a consequence, each geochronologic era boundary in the Precambrian prior to the base of the Ediacaran is defined by a Global Standard Stratigraphic Ages (GSSA), which is a numeric time coordinate.

In an earlier paper (Cox and Richard 2005), we formalized the static relationships involved in the definition of a geologic timescale as a Unified Modeling language (UML) (ISO/IEC 2005) class model, referred to hereinafter as C&R. This was based on ICS principles, and was also linked to the schema for temporal objects and reference systems described in ISO 19108 (ISO/TC-211 2002), to the model for Observations and Measurements described in ISO 19156 (ISO/TC-211 2011), and integrated with the GeoSciML model for geologic information from the Commission for Geoscience Information (CGI) (Richard and CGI-Interoperability-Working-Group 2007). The CHRONOS project provided services using the XML serialization of the timescale from GeoSciML (Fils et al. 2009).

A number of RDF-based geologic timescales and services have been described. The SWEET system of ontologies for the earth and environmental sciences includes a geologic timescale that uses a basic temporal topology in which each time unit is described using a start and end expressed as a numeric time instant (Raskin and Pan 2005), and thus provides a snapshot of the calibrated timescale but without the

[1] http://www.stratigraphy.org

GSSP logic. A SKOS-based system was developed by (Ma et al. 2011), with a particular focus on multi-lingual support, but limited formalization of the timescale logic beyond what is provided generically by SKOS thesaurus relations. (Perrin et al. 2011) developed an ontology focussed on stratigraphic correlations, in which the core model was inspired by C&R, but is aligned with the W3C Time ontology (Hobbs and Pan 2006). The timescale ontology developed by Ma (Ma et al. 2012) focuses on the use of the hierarchical relationships in the timescale to annotate and query geologic databases. Ma and Fox (2013) provide a comprehensive review of these and other prior work, and we also draw attention to some limitations of the previous work later in this paper.

In this paper we describe an OWL2 ontology (W3C OWL Working Group 2012) derived from the C&R model, and a set of vocabulary instances formalized using this ontology, whose content is based on versions of the timescale published by the ICS as the International [Chrono]stratigraphic Chart (ICC). The intention was to develop rich representations, using contemporary semantic web technology, aligned both with stratigraphic best practice and with standard treatments of temporal systems from ISO. The vocabularies have been published using a number of web interfaces and APIs, which support their use both as a reference system by geologists, and for investigations concerning the timescale itself, its evolution, and its reference material.

The paper is structured as follows: in section 2. we review the C&R model for the timescale and its relationship with observed geology; section 3. has a brief description of the UML→OWL conversion rules; section 4. systematically describes the axioms of the timescale ontology, and alignment with some other ontologies; in section 5. we describe the implementation of instances of the geological timescale based on the ICC, and publication through various web APIs; in section 6. we discuss how the approach to versioning used aligns with stratigraphic practice, and make some comparisons with previous attempts to formalize the geologic timescale in a modern machine-processable form. We conclude in section 7.

2. Review of C&R geologic timescale model

2.1 Reference systems

Conventional historical geology is based on the classification of the age of geological units and other features using terms from the geological timescale. This is an example of an **ordinal reference system**, in which the goal is to support classification and ordering of items in scope using an ordered set of named categories. In contrast, ages specified chronometrically use a **coordinate system**, which measures distance, in appropriate units, in a specified direction from some datum or origin. Other common types of reference system are a **nominal reference system**, which assigns a named classifier but with no ordering significance, and a **measure**, which is based on absolute amount expressed as a scaled number.

While numeric ages have become more common and more important as chronometric methods have developed and become more easily available, geologic datasets and the geologic literature are full of descriptions that use the traditional chronostratigraphic ages, with the expectation that geologically literate readers understand the significance of these, including their ordering relationships.

2.2 ISO 19108 Temporal Schema

2.2.1 Temporal Reference Systems

ISO 19108 (ISO/TC-211 2002) provides a comprehensive model of temporal classes and reference systems, whose scope extends significantly beyond the familiar calendar/clock based systems. Four types of temporal reference system are defined (Figure 1):

- **Clock** - supporting a time position given as a time-of-day, usually in hours-minutes-seconds;
- **Calendar** - allowing a time position to be given using a calendar system, usually as year-month-day;
- **Temporal Coordinate System** - which allows a time position to be given as a number on a scale with a datum and direction (e.g. seconds counting forward from a specific date and time; years counting backwards from a specific date);
- **Temporal Ordinal Reference System** - which is an ordered system of named intervals.

As noted in C&R, in the ISO 19108 Temporal Ordinal Reference System the values of both the begin and end of a TM_OrdinalEra class are of type DateTime, so this structure may only be used directly for systems defined using calendar/clock frames.

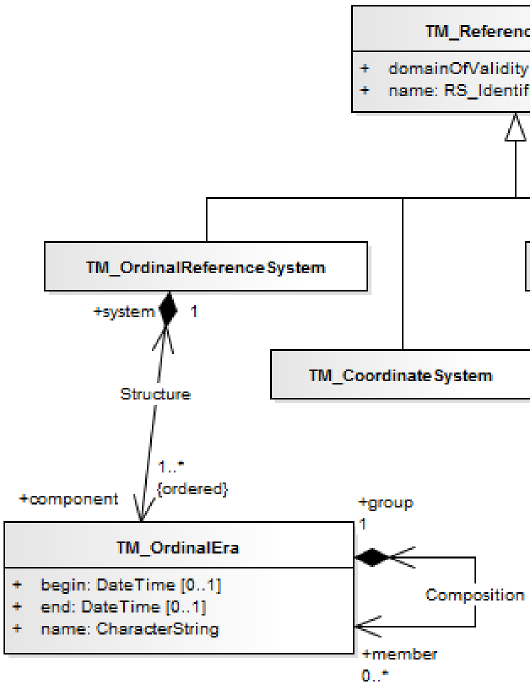

Figure 1: UML structure diagram of temporal reference systems from ISO 19108:2002, with details of temporal ordinal reference system

Note that a TM_OrdinalEra may be composed of subordinate member ordinal eras, defining a hierarchy. There is an implicit logical constraint that the time interval represented by a member must be completely within the time interval represented by the containing group ordinal era.

2.2.2 Temporal Objects

For temporal objects, ISO 19108 distinguishes between (Figure 2):

- temporal *geometry,* which is composed of time instants and periods given as values relative to one of the reference systems, and
- temporal *topology*, which is composed of time nodes and edges, where the temporal objects are defined in terms of their ordering relationships.

In the temporal topology model the start or end of a temporal edge is a temporal node. A temporal topological primitive is aligned to a temporal reference system through a *geometry* property, which links to the corresponding temporal geometric primitive: edge to period and node to instant.

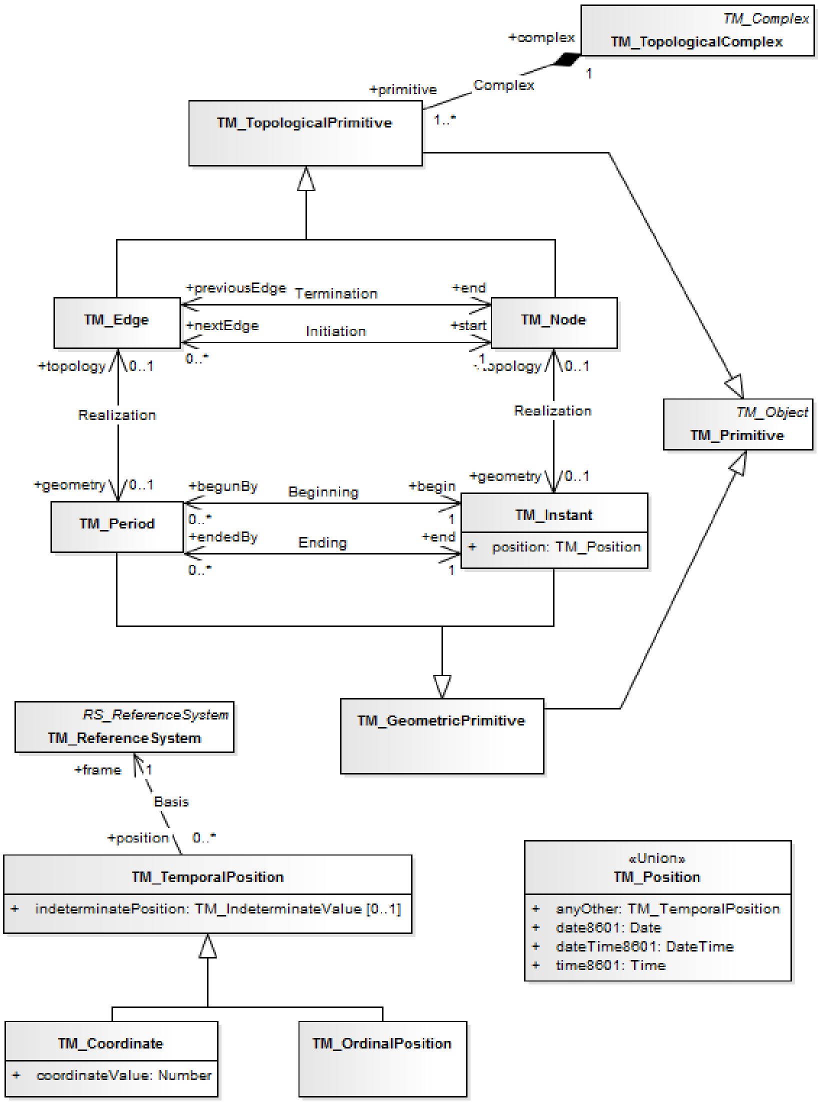

Figure 2: Temporal primitives from ISO 19108:2002.

2.3 C&R model for geologic timescale

2.3.1 Temporal Hierarchical Ordinal Reference System

In order to support the requirements of the geological timescale, Cox and Richard (2005) defined a new temporal reference system known as THORS (Temporal Hierarchical Ordinal Reference System) in which the era start and end is modeled as a class in its own right, rather than as a simple date/time attribute (Figure 3). C&R also discussed the way that a temporal ordinal reference system could be modeled as a specialized temporal topology, with a constraint on the permissible subdivision of edges at each rank level. However, because of the need to avoid multiple-inheritance (for XML implementation reasons), while THORS follows the pattern for temporal topology, it is not formally linked to the ISO 19108 temporal topology models.

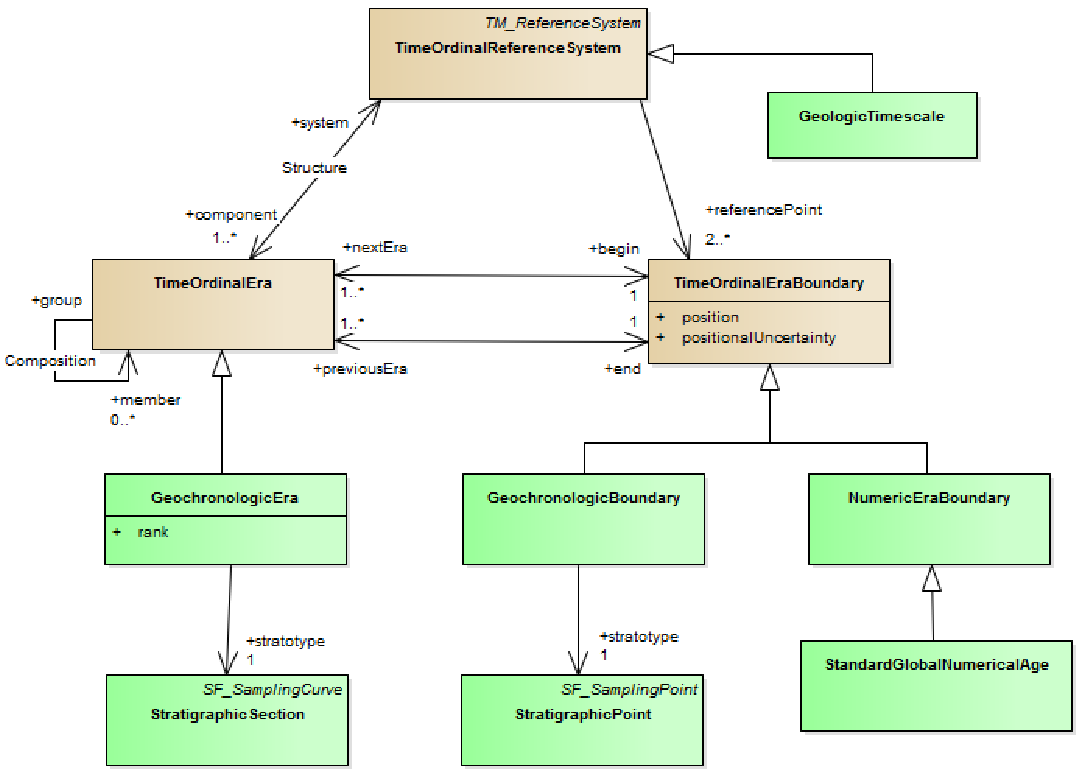

Figure 3: Time ordinal reference system (brown) and geological timescale model (green) [adapted from Cox & Richard 2005].

The elements of the geologic timescale model (GTS) are specialized from the THORS classes as shown in FIgure 3. We retained the term ‘era’ in the name GeochronologicEra for consistency with the base model from ISO 19108, notwithstanding the fact that ‘Era’ is used in the geologic timescale to denote a specific rank of interval in the geologic timescale. Two kinds of boundaries are recognised:

GeochronologicBoundary which is defined using stratigraphic principles, and NumericEraBoundary whose position is fixed conventionally at specified dates.

2.3.2 Global Stratotype Section and Point

C&R also modeled the links from the timescale to stratotypes, the stratigraphic sections and points, that anchor boundaries (Figure 3) and the relationships from these to geologic features from the GeoSciML model (Figure 4). Following the practice defined by the ICS, a GeochronologicBoundary is linked with a StratigraphicPoint stratotype that anchors the chronologic position of the boundary in the material geologic record formed at that time. A GeochronologicEra may be linked to a corresponding StratigraphicSection stratotype, comprising rocks that formed during that time interval. If ratified by the ICS as a Global Stratotype Section and Point (GSSP) then a StratigraphicPoint stratotype is known informally as a ‘Golden Spike’, and this symbol appears on the graphical representation of the timescale (see http://stratigraphy.org/index.php/ics-chart-timescale). In the C&R model, stratotypes are defined as a specialization of the appropriate SamplingFeature from the ISO 19156 (ISO/TC-211 2011) model (Figure 4). Stratotypes thus have geometry properties and relationships to other sampling features as defined by ISO 19156, with additional properties as required by the ICS model. Features from GeoSciML (Richard and CGI-Interoperability-Working-Group 2007) may be linked to the timescale elements, as indicated by the incoming relationships from Contact and GeologicEvent classes.

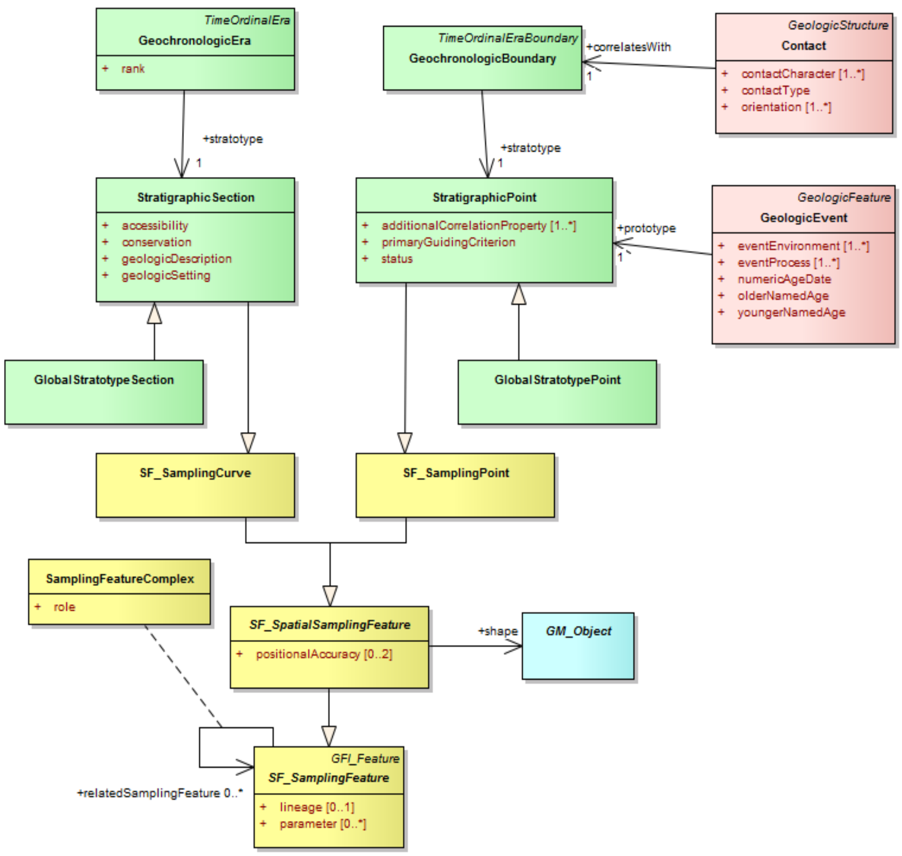

Figure 4: Relationships of stratigraphic points and sections from GTS model (green) with ISO 19156 Sampling feature model (yellow), GeoSciML (pink) and ISO geometry (blue) (adapted from Cox & Richard, 2005).

3. ISO UML→ OWL conversion rules

The geologic timescale ontology is derived from C&R largely by application of conversion rules for ISO-formalized models. Project 19150-2 in ISO/TC 211 is developing rules for transforming the UML models that define the components of the ISO/TC211 'Harmonized Model' into OWL ontologies, as well as rules for developing Application Schemas directly in OWL. This is currently in Draft International Standard status (ISO/TC-211 2014a). The Harmonized Model is a set of standard models for geometry, temporal, metadata, spatial fields, coordinate systems, etc, which may be used in the development of geographic 'Application Schemas', i.e. domain- and application-specific information models. The model is formalized using a UML profile described in ISO 19103 (ISO/TC-211 2005a) and ISO 19109 (ISO/TC-211 2005b)[2], in which UML is used as a "conceptual" modeling language, where the classes are generally intended to represent phenomena in the real world, and not merely the documents describing them. This approach may be contrasted with initiatives that provide a UML profile for graphical representation of RDF and OWL ontologies (e.g. Object Management Group's ODM (Object Management Group 2009)).

The ISO 19150-2 rules define the mappings and some URI patterns:

UML package → OWL ontology (and RDF namespace)
UML class → OWL class, with specialization modeled as sub-classing
UML attribute and association-role → RDF Property.

The UML stereotypes from the profile are implemented as OWL Classes, so that classes that would have this stereotype in UML are implemented as sub-classes of the corresponding OWL class (except for 'Enumeration' for which the natural implementation is a rdfs:Datatype).

A partial conversion of the ISO Harmonized Model (Cox 2013) is available at http://def.seegrid.csiro.au/isotc211/ containing representations of all the dependencies of the geologic timescale ontology. OWL Ontologies corresponding to ISO 19108 Temporal Schema and ISO 19156 Sampling Features are provided[3]. The definition of temporal edge within this ontology, shown in Listing 1, illustrates the general style.

Listing 1: OWL definition of Temporal Edge (Turtle format, Beckett & Berners-Lee [2011]). The property named 'start' in the UML is changed to 'begin' in OWL for consistency between topological and geometric cases.

```
tm:Edge
      a       owl:Class ;
      rdfs:isDefinedBy <http://def.seegrid.csiro.au/isotc211/iso19108/2002/temporal> ;
      rdfs:label "Temporal Edge"^^xsd:string ;
      rdfs:subClassOf tm:TopologicalPrimitive ;
      rdfs:subClassOf [ a       owl:Restriction ;
                owl:cardinality "1"^^xsd:nonNegativeInteger ;    owl:onProperty tm:end    ] ;
      rdfs:subClassOf [ a       owl:Restriction ;
                owl:allValuesFrom tm:Node ;                      owl:onProperty tm:begin   ] ;
```

2 These standards are currently in revision, in part to improve alignment with OWL implementations, see ISO/DIS 19103 (ISO/TC-211 2014b) and ISO/DIS 19109 (ISO/TC-211 2013).

3 http://def.seegrid.csiro.au/isotc211/iso19108/2002/temporal, http://def.seegrid.csiro.au/isotc211/iso19156/2011/sampling

```
rdfs:subClassOf [ a       owl:Restriction ;
          owl:cardinality "1"^^xsd:nonNegativeInteger ;    owl:onProperty tm:begin  ] ;
rdfs:subClassOf [ a       owl:Restriction ;
          owl:maxCardinality "1"^^xsd:nonNegativeInteger ; owl:onProperty tm:geometry  ] ;
rdfs:subClassOf [ a       owl:Restriction ;
          owl:allValuesFrom tm:Period ;                    owl:onProperty tm:geometry  ] ;
rdfs:subClassOf [ a       owl:Restriction ;
          owl:allValuesFrom tm:Node ;                      owl:onProperty tm:end  ] ;
skos:notation "TM_Edge"^^h2o:ISOClassName ;
skos:prefLabel "Temporal edge"@en .
```

These conversion rules and standard components provide the basis for the implementation of the C&R model as an ontology, as described in the next section.

4. The geologic timescale ontology

4.1 Namespaces

The principal RDF namespaces and corresponding prefixes used in the OWL implementation of the geologic timescale model are given in Table 1. Note that the classes and properties in the ontologies are denoted with 'hash URIs' which mean that they are accessed in the context of the complete ontology, when using HTTP (Berrueta and Phipps 2008).

Table 1: Namespaces used in timescale ontology

RDF namespace	Prefix	Source
http://def.seegrid.csiro.au/isotc211/iso19103/2005/basic#	basic:	ISO 19103:2005; ISO/DIS 19150-2; Cox, 2013a
http://resource.geosciml.org/ontology/timescale/gts#	gts:	This paper
http://def.seegrid.csiro.au/isotc211/iso19156/2011/sampling#	sam:	ISO 19156:2011; ISO/DIS 19150-2; Cox, 2013a
http://www.w3.org/2004/02/skos/core#	skos:	Miles & Bechhofer, 2009
http://resource.geosciml.org/ontology/timescale/thors#	thors:	This paper
http://def.seegrid.csiro.au/isotc211/iso19108/2002/temporal#	tm:	ISO 19108:2002; ISO/DIS 19150-2; Cox, 2013a

4.2 THORS Ontology

The Temporal Hierarchical Ordinal Reference System (THORS) ontology implements a temporal ordinal reference system that combines the ISO 19108 temporal topology model (TT) and temporal ordinal reference system model (TORS). The key structure is shown in Figure 5.

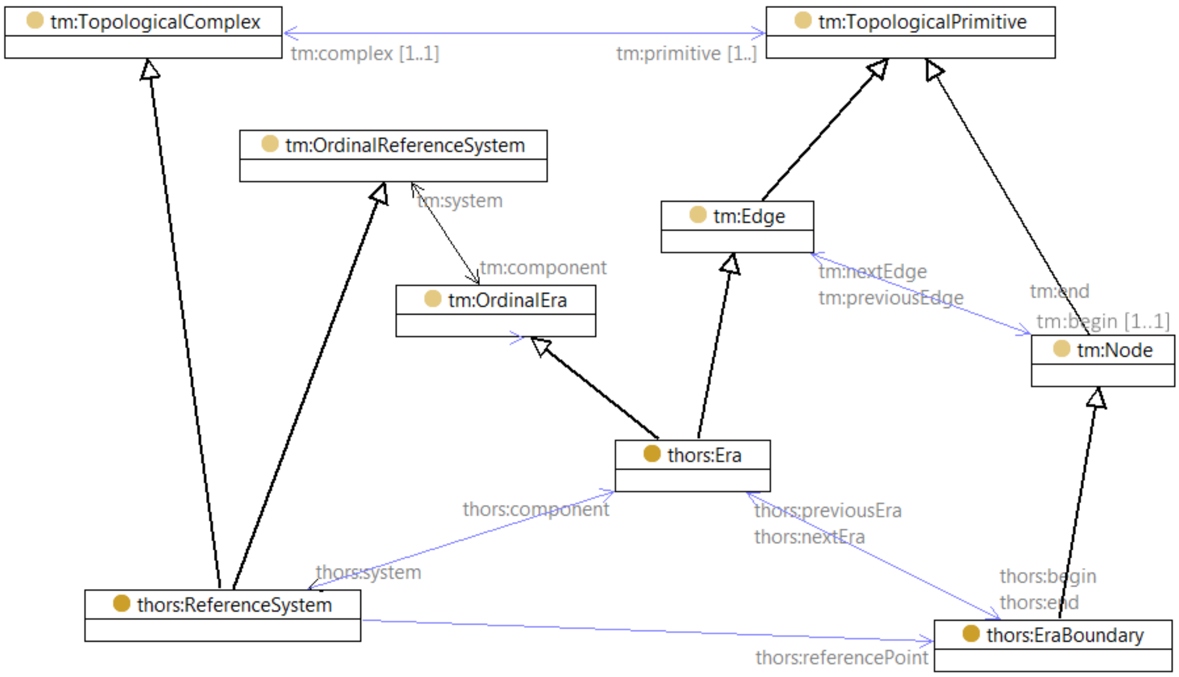

Figure 5: Classes in the temporal hierarchical ordinal reference system ontology (THORS) (UML-style diagram of OWL classes and their relationships)

Three classes are defined (Table 2). Class names are simplified within THORS since there is no ambiguity with other reference system types, unlike in the ISO 19108 model. Furthermore, in contrast to the UML-XSD pipeline, in OWL/RDF there is no practical impediment to multiple inheritance if this captures the model correctly, so thors:ReferenceSystem and thors:Era are subclassed from classes in the OWL implementations of both temporal topology and temporal ordinal reference system. Though not shown in the diagram, the THORS classes are also linked to SKOS classes, to enable a timescale instance to be compatible with generic vocabulary APIs based on SKOS (Cox et al. 2014).

Table 2: Classes in the THORS ontology

Class	subClassOf
thors:ReferenceSystem	tm:TopologicalComplex , tm:OrdinalReferenceSystem , skos:ConceptScheme
thors:Era	tm:Edge , tm:OrdinalEra , skos:Concept
thors:EraBoundary	tm:Node , skos:Concept

Nine properties are defined (Table 3). A key difference between frame-based models (UML, XSD) and OWL is the support for property specialization, using the RDFS subPropertyOf property. All of the THORS properties are defined as sub-properties of properties defined in the ISO 19108 TT and TORS models, as well as appropriate SKOS properties. This links the THORS ontology to the full semantics of the parent ontologies.

Table 3: Properties in the THORS ontology

property	**subPropertyOf**	**domain**	**range**	**inverse**
thors:component	tm:component tm:primitive skos:hasTopConcept	thors:ReferenceSystem	thors:Era	thors:system
thors:referencePoint	tm:primitive skos:hasTopConcept	thors:ReferenceSystem	thors:EraBoundary	
thors:system	tm:complex tm:system skos:topConceptOf	thors:Era	thors:ReferenceSystem	thors:component
thors:member	tm:member tm:primitive skos:narrower	thors:Era	thors:Era	
thors:begin	tm:begin skos:semanticRelation	thors:Era	thors:EraBoundary	thors:nextEra
thors:end	tm:end skos:semanticRelation	thors:Era	thors:EraBoundary	thors:previousEra
thors:nextEra	tm:nextEdge skos:semanticRelation	thors:EraBoundary	thors:Era	thors:begin
thors:previousEra	tm:previousEdge skos:semanticRelation	thors:EraBoundary	thors:Era	thors:end
thors:positionalUncertainty	basic:measure	tm:TemporalPosition	basic:Time	

4.3 GTS Ontology

The GTS ontology defines a structure for geologic timescales derived from the THORS, and also linked to geology classes used in calibration, according to the ICS procedures. The key structure is shown in Figure 6.

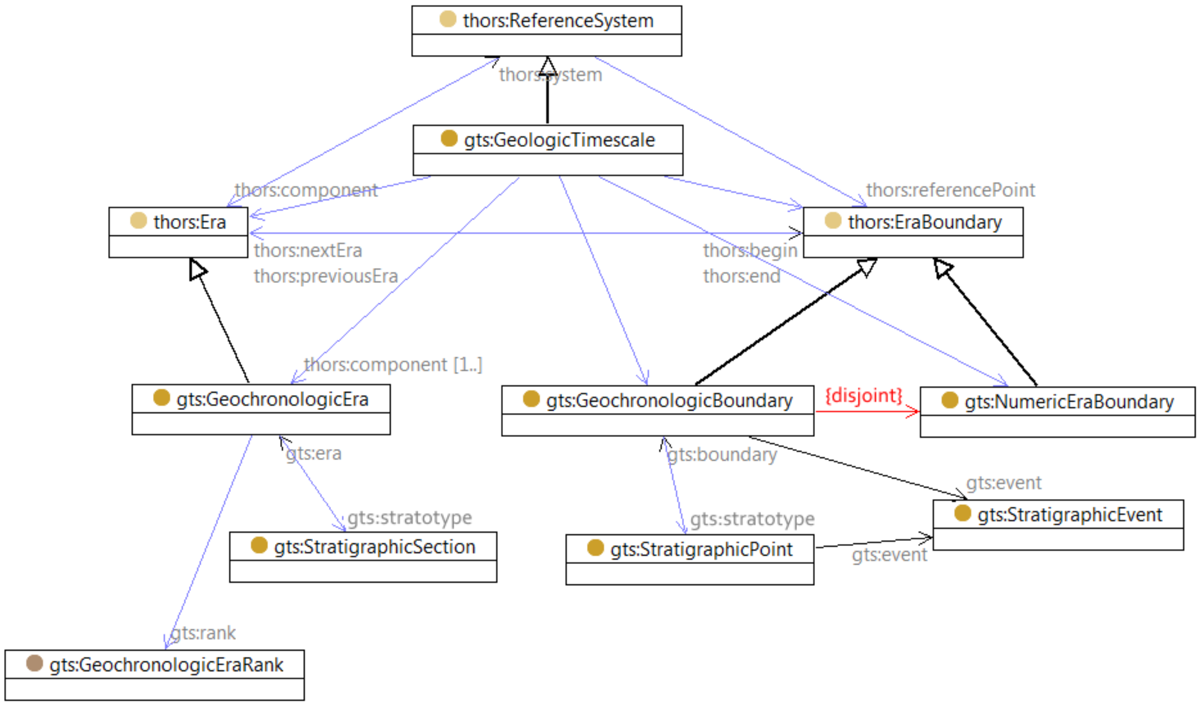

Figure 6: Key classes in the geologic timescale ontology (GTS)

Eight classes are defined, subclassed from the appropriate class from THORS, from ISO 19156 (ISO/TC-211 2011) (sampling features) ontologies, or from SKOS (Miles and Bechhofer 2009) (Table 4).

Table 4: Classes in the GTS ontology

Class	subClassOf
gts:GeologicTimescale	thors:ReferenceSystem
gts:GeochronologicEra	thors:Era
gts:GeochronologicEraRank	skos:Concept
gts:GeochronologicBoundary	thors:EraBoundary
gts:NumericEraBoundary	thors:EraBoundary
gts:StratigraphicSection	sam:SamplingCurve
gts:StratigraphicPoint	sam:SamplingPoint
gts:StratigraphicEvent	-

13 properties are defined in the GTS ontology (Table 5). In addition to properties specialized from the THORS model, GTS adds the properties required to capture elements of the geologic timescale and its calibration using GSSPs as shown in Figure 4, and some inverse properties (gts:boundary, gts:era, gts:manifestedBy, gts:event). The property gts:boundaryLevel provides a more precise slot to record the position of the stratotype, and the status of the stratotype is recorded with two properties (a boolean flag gts:ratifiedGSSP, and a status description).

Table 5: Properties in the GTS ontology

property	domain	range
gts:boundary	gts:StratigraphicPoint	gts:GeochronologicBoundary
gts:era	gts:StratigraphicSection	gts:GeochronologicEra
gts:rank	gts:GeochronologicEra	gts:GeochronologicEraRank
gts:stratotype	gts:GeochronologicBoundary or gts:GeochronologicEra	gts:StratigraphicPoint or gts:StratigraphicSection
gts:manifestedBy inverseOf gts:correlatesWith	gts:GeochronologicBoundary or gts:GeochronologicEra	(Geologic Feature)
gts:correlatesWith inverseOf gts:manifestedBy	(GeologicFeature)	gts:GeochronologicBoundary or gts:GeochronologicEra
gts:event	gts:GeochronologicBoundary or gts:GeochronologicEra	gts:StratigraphicEvent
gts:status	gts:NumericEraBoundary or gts:StratigraphicPoint	(statement about ICS ratification status)
gts:ratifiedGSSP	gts:StratigraphicPoint or gts:GeochronologicBoundary or gts:NumericEraBoundary	boolean
gts:correlationEvent (*primaryGuidingCriterion* in the UML version)	gts:StratigraphicPoint	(description of the geologic event that is the basis for boundary)
gts:boundaryLevel	gts:StratigraphicPoint	(description of the location of the calibration point within the geologic section)
gts:geologicDescription	gts:StratigraphicSection	(description of the geologic section)
gts:geologicSetting	gts:StratigraphicSection	(description of the overall geologic setting of the section and point)

The term ‘era’ in the name GeochronologicEra matches the generic use reflected in ISO 19108, notwithstanding the fact that ‘Era’ is used in the geologic timescale to denote a specific rank of interval in the geologic timescale. The full set of rank values used in the International Chronostratigraphic Chart is included in GTS[4].

5. The International Chronostratigraphic Chart

5.1 Publication and versioning

The International Chronostratigraphic Chart (and its predecessors, known as The International Stratigraphic Chart) is published by the International Commission for Stratigraphy on a semi-periodic schedule. The contents are provided online in the following forms:

1. a table of the current set of GSSPs[5]
2. an image, in either PDF or PNG format[6].

The image representation summarizes the hierarchical arrangement of the geochronologic units, the calibrated age of each boundary where known, and also indicates which boundaries have a Global Stratotype Point ratified by the International Commission for Stratigraphy (a ‘golden spike’).

[4] See http://resource.geosciml.org/ontology/timescale/rank which includes Super-Eon , Eon , Era , Period , Sub-Period , Epoch , Age.
[5] http://stratigraphy.org/GSSP/index.html
[6] http://www.stratigraphy.org/index.php/ics-chart-timescale

Significant versions were published in 2004, 2005, 2006, 2008, 2009, 2010, 2012, 2013, 2014. A full history of the image representations is available from the GeoSciML Wiki site[7]. Comparison of different versions shows changes to the timescale in four aspects:

1. era names
   e.g. 2010 'Ionian' → 2012 Middle Pleistocene,
   2010 Cambrian Stage 9 → 2012 Jiangshanian
2. era hierarchy
   e.g. Gelasian: 2006 part of Pliocene → 2009 part of Pleistocene,
   Quaternary: 2006 informal → 2009 formal Period
3. boundary ages and/or uncertainty
   e.g. Base of Pennsylvanian: 2010 318.1 ±1.3 Ma → 2012 323.2 ±0.4 Ma
4. ratification of boundaries
   e.g. 2012 Base of Lutetian, 2012 Base of Jiangshanian

5.2 OWL implementation

We have prepared a separate OWL implementation of the International Chronostratigraphic Chart for each version published since 2004 (Cox and Richard 2014). The namespace for all individuals in the charts is http://resource.geosciml.org/classifier/ics/ischart/. The prefix **isc**: is used in compact URIs, with the following patterns:

Geochronologic era	isc:{era name}
Geochronologic boundary	isc:Base{era name}
Temporal position	isc:Base{era name}Time
Stratigraphic point	isc:[GS]SPBase{era name}
Timescale	isc:{publication year}

Individual elements in the timescales are denoted with 'slash URIs' which allow them to be accessed individually using HTTP (Berrueta and Phipps 2008). Listing 2 shows a sample description of a geochronologic era and boundary, and associated elements. Although the model and ontology does support it, the current dataset does not include descriptions of the Stratigraphic Sections as this is not easily obtainable from the www.stratigraphy.org website.

Listing 2: Definition of Middle Jurassic Epoch - Turtle format (Beckett et al. 2014)

```
isc:MiddleJurassic
      a       gts:GeochronologicEra , skos:Concept ;
      rdfs:comment "older bound-175.6 +/-2"@en , "younger bound-161.2 +/-4"@en ;
      rdfs:label "Middle Jurassic Epoch"@en ;
      gts:rank <http://resource.geosciml.org/ontology/timescale/rank/Epoch> ;
      thors:begin isc:BaseMiddleJurassic ;
      thors:end isc:BaseUpperJurassic ;
      thors:member isc:Bajocian , isc:Aalenian , isc:Callovian , isc:Bathonian ;
      =       <http://dbpedia.org/resource/MiddleJurassic> ;
      skos:altLabel "中侏罗统"@zh , "ジュラ紀中期"@ja ;
      skos:broader isc:Jurassic ;
```

[7] https://www.seegrid.csiro.au/wiki/CGIModel/GeologicTime#ICS_Stratigraphic_Chart

```
      skos:broaderTransitive
              isc:Mesozoic , isc:Phanerozoic , isc:Jurassic ;
      skos:inScheme <http://resource.geosciml.org/classifier/ics/ischart/2013> ;
      skos:narrower isc:Bajocian , isc:Aalenian , isc:Callovian , isc:Bathonian ;
      skos:narrowerTransitive
              isc:Bajocian , isc:Aalenian , isc:Callovian , isc:Bathonian ;
      skos:notation "a1.1.2.2.2"^^gts:EraCode ;
      skos:prefLabel "Mittlerer Jura"@de , "中侏罗世"@zh , "giurassico medio"@it , "Jurassique moyen"@fr ,
"Middle Jurassic"@en , "srednja jura"@sl , "Jurássico Médio"@pt , "Keski-Jura"@fi , "középső-jura"@hu ,
"Kesk-Juura"@et , "Mellem Jurassisk"@da , "Средна Юра"@bg , "Midtre jura"@no , "Midden Jura"@nl , "中期ジュ
ラ紀"@ja , "Jurásico Medio"@es , "Střední jura"@cs , "stredná jura"@sk , "Vidurinė Jura"@lt , "mellersta
jura"@sv , "Środkowa Jura"@pl ;
      foaf:isPrimaryTopicOf
              <http://sweet.jpl.nasa.gov/2.2/stateTimeGeologic.owl#MiddleJurassic> .

isc:BaseMiddleJurassic
      a       gts:GeochronologicBoundary , skos:Concept ;
      rdfs:label "Base of Middle Jurassic"@en ;
      tm:temporalPosition isc:BaseMiddleJurassicTime ;
      gts:stratotype isc:GSSPBaseMiddleJurassic ;
      thors:nextEra isc:Aalenian , isc:MiddleJurassic ;
      thors:previousEra isc:LowerJurassic , isc:Toarcian ;
      skos:altLabel "Base of Aalenian"@en ;
      skos:inScheme <http://resource.geosciml.org/classifier/ics/ischart/2013> ;
      skos:prefLabel "Base of Middle Jurassic"@en ;
      skos:topConceptOf <http://resource.geosciml.org/classifier/ics/ischart/2013> .

isc:BaseMiddleJurassicTime
      a       tm:Coordinate ;
      tm:frame <http://resource.geosciml.org/classifier/cgi/geologicage/ma> ;
      tm:value "174.1"^^xsd:float ;
      thors:positionalUncertainty
              [ a       basic:Time ;
                basic:value "1.0"^^xsd:float ;
                basic:uom <http://www.opengis.net/def/uom/UCUM/0/Ma>
              ] .

isc:GSSPBaseMiddleJurassic
      a       gts:StratigraphicPoint ;
      rdfs:comment "Fuentelsaz, Spain " ;
      rdfs:label "GSSP of Base Middle Jurassic"@en ;
      rdfs:seeAlso "<http://www.stratigraphy.org/GSSP/Aalenian.pdf>" ;
      sam:shape
              [ a       gm:Point ;
                gm:position
                        [ a       gm:Position ;
                          gm:coordinates "41.1708 -1.8333"^^basic:ordinates ;
                          gm:srs  <http://www.opengis.net/def/crs/EPSG/0/4326>
                        ]
              ] ;
      dc:source "Episodes 24/3, p.166 175, 2001"^^xsd:string ;
      gts:boundaryLevel "base of Bed FZ 107 "^^xsd:string ;
      gts:correlationEvent
              "Ammonite FAD Leioceras opalinum and Leioceras lineatum" ;
      gts:ratifiedGSSP "true"^^xsd:boolean ;
      gts:status "Ratified 2000"@en .
```

In addition to the structural elements of the timescale, which match the ontology described above, the following data is included:

1. each geochronologic era name is provided in multiple languages, using skos:prefLabel
2. certain class and property generalizations implied by the ontology are included as assertions in the vocabulary: specifically skos:Concept, skos:narrower
3. some inverse and transitive properties are included as assertions: skos:broader, skos:broaderTransitive, skos:narrowerTransitive
4. mappings to eras in the DBPedia (Bizer et al. 2009) and SWEET (Raskin and Pan 2005) versions of the geologic timescale are included
5. for each GSSP the original publication is provided, using a text citation in dc:source, and a link to a copy of the article on the www.stratigraphy.org website using rdfs:seeAlso. .

5.3 Publication

Each version of the OWL timescale is published in the following ways:

1. as a static file at `http://resource.geosciml.org/vocabulary/timescale/isc{year}`
2. through a SPARQL endpoint at `http://resource.geosciml.org/sparql/isc{year}`
3. at a SISSvoc service at `http://auscope-services.arrc.csiro.au/sissvoc/isc{year}/collection`

SISSvoc is a generic vocabulary search interface, which relies on classes and properties from the SKOS vocabulary. While SISSvoc uses SKOS elements for searching, when it reports the description of a concept, it includes all triples from any vocabulary related to the resources of interest, so may be used for an arbitrary vocabulary provided it is at least 'decorated' with SKOS-related axioms. The links to SKOS included in the THORS and GTS ontologies allow axioms involving SKOS classes and predicates to be inferred in the timescale instances, as shown above. Thus, while the timescale instances are structured according to the chronostratigraphic logic, the published vocabularies also contain SKOS types and properties.

SISSvoc provides a generic HTML user-interface to a browser, and alternative RDF, XML, text and JSON through content negotiation using HTTP Accept headers. The service hosting the actual URIs for individual elements of the timescale, such as `http://resource.geosciml.org/classifier/ics/ischart/Pliocene`, automatically redirects requests to a representation provided by the most recent SISSvoc service. At time of writing this is the 2014 version, so the URI above is redirected to
`http://auscope-services.arrc.csiro.au/sissvoc/isc2014/resource`
`?uri=http://resource.geosciml.org/classifier/ics/ischart/Pliocene`
which can be read as "what isc2014 knows about Pliocene". The date-stamped SISSvoc service provides the context for a graph of statements about a concept from the geologic timescale, at the corresponding publication date.

In addition, the full set of timescales is bundled in a compressed file and is available from the CSIRO data repository (Cox and Richard 2014).

6. Discussion

6.1 Stratigraphic practice and timescale versions

To assign an age to a geological object using stratigraphic methods, a geologist seeks features that can be observed in the object that are characteristic of a particular age. Various kinds of feature are used, the requirement being that they show distinctive patterns or changes that can be correlated between observations in different localities. Quintessential correlatable features are the emergence or disappearance of fossil instances of evolving taxa, but chemical, magnetic, structural and even lithological characteristics are also used. The GSSP program is based on agreement upon a standard global correlation event for each boundary, together with a locality where the evidence of this event can be observed in a continuous sequence of rocks formed before, during, and after the boundary event (Murphy and Salvador 1994; Remane et al. 1996). The challenge is to identify short-duration events that have left observable manifestations in the rock record over a wide area, ideally globally. Large bolide impacts, such as that related to the Cretaceous-Tertiary (K-T) boundary, are an example of such events. Secondary events and their evidence may then be correlated with the primary event in order to enable correlation between rocks formed in a variety of settings that may not support observable evidence of the primary event. The disappearance of dinosaurs is such a secondary event correlated to the K-T bolide impact.

For ordinal eras that are defined based on GSSPs, a change in the numeric calibration of the timescale does not affect previous assignment of a body of material to a named era, if the assignment was based on characteristics in the rock not tied directly to the absolute age. For example, the beginning of the Phanerozoic Eon is marked by the widespread appearance in the Cambrian Period of fossils of certain types of multicellular organisms whose predecessors are essentially unknown. The base of the Phanerozoic had been believed to be around 570 Ma, but in 1994 when the GSSP was assigned, this was revised to around 540 Ma. Nevertheless, rocks classified as Cambrian in reports from the 1960s on the basis of their fossil content are still correlated with Cambrian rocks in reports from the 2000s, even though the understanding of the absolute (chronometric) age of the Cambrian Period has changed. Generally, the intention of practitioners is that the term ‘Cambrian’ denotes a time interval represented in the geologic record by rocks with particular characteristic fossils (Shergold and Cooper 2005). The temporal coordinate position may change as more information about the boundaries becomes available.

In our implementation, except for the timescale as a whole, the URI denoting a timescale element includes no date or other version indicator. Each era and boundary has the same identifier in all versions of the timescale in which it appears. This matches the naming practice of the ICS, and the stratigraphic correlation principles described above, where the intention is that geochronologic eras are conceptually invariant, though what we know about them may change with time, and be realized as different RDF graphs containing the same nodes. In a dataset or report, the age of a geologic unit or other feature may be classified using an era from the geologic timescale, denoted by its URI, which is independent of the version of the timescale. This matches stratigraphic practice, as the geochronologic classification remains the same, though the description of that classifier may change subsequent to the time of assignment. The most recent version of the timescale contains updated descriptions of the concepts that were present in earlier versions.

For the part of geologic history in which the boundaries of the timescale eras are GSSAs, broad scale correlation relies on isotopically-determined ages. Since the eras are defined chronometrically, a change to the numeric value of the boundary would fundamentally redefine the era, so a change in identifier should be required. However, the only significant change in the datasets under consideration was the insertion of the Hadean Eon with the base of the Archean fixed at 4000 Ma between the 2006 and 2008 versions of the chart. Since there is no geologic record for the Hadean, this means that for practical purposes the definition of the Archean (i.e. the era containing the earliest rocks) did not change.

6.2 Formalization of the Quaternary changed the timescale topology

Probably the most radical changes in the official timescale over the period considered are related to the formalization of the Quaternary Period between 2006 and 2009, and an associated re-assignment of the Gelasian Age from the top of the Pliocene Epoch to the bottom of the Pleistocene Epoch. This is illustrated in directed graphs from part of the 2006 and 2009 versions of the timescale shown in Figure 7, showing how the formalization of the timescale is realized in the RDF representation.

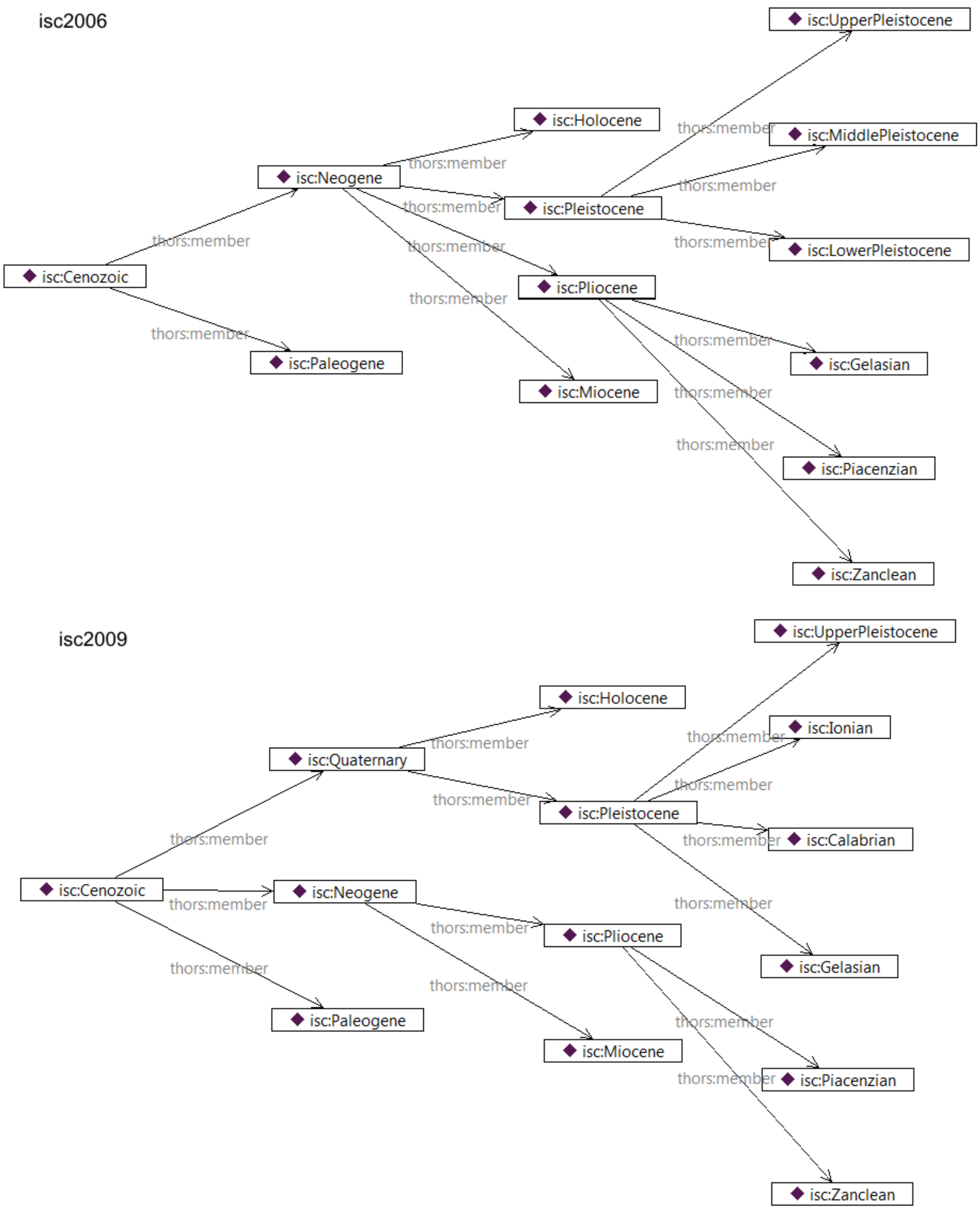

Figure 7: Topology of Cenozoic, as defined in the 2006 and 2009 versions of the ICS timescale

6.3 Comparison with previous approaches

In the introduction we mentioned a number of prior attempts to formalize the geologic time.

An XML Schema representation of the Cox and Richard (2005) model was provided in GeoSciML 3.0 (Richard and CGI-Interoperability-Working-Group 2007), and used by Fils et al. (2009). This was an effective serialization of the model, but only geospatial software that can process the GML-based style of XML can use it. Since the timescale is a reference-system it is best delivered in a form with direct links to the individual concepts. Standard tooling in semantic web technology supports this approach.

SWEET (Raskin and Pan 2005) provides a compact representation of the timescale in OWL. However, the only information provided for the beginning and end of each era is an age, and there is no other information about the boundaries of temporal eras. In particular, there is no link to prototype occurrences used to calibrate the timescale, which are the focus of the standardization activity of the International Commission for Stratigraphy.

Furthermore, in the SWEET ontologies, the base URI for all concepts is changed with every version of SWEET. Thus the Pliocene Epoch has been identified variously as

```
http://sweet.jpl.nasa.gov/1.1/time.owl#PLIOCENE
http://sweet.jpl.nasa.gov/2.0/timeGeologic.owl#Pliocene
http://sweet.jpl.nasa.gov/2.1/reprTimeGeologicPeriod.owl#Pliocene
http://sweet.jpl.nasa.gov/2.2/stateTimeGeologic.owl#Pliocene
http://sweet.jpl.nasa.gov/2.3/stateTimeGeologic.owll#Pliocene
```

A human can infer relationships between these concepts from inspecting the identifiers. However, to any RDF- or OWL-aware processor these are different URIs for different resources, and since there are no explicit relationships from later to earlier versions, there is no way to do any reasoning involving them. Effectively, each URI identifies a description or graph contained in the specific version of SWEET, which also does not correspond to versioning by the governing organization for the geologic timescale. In contrast, in the vocabularies described in this paper, the URI for an element of the timescale identifies the concept, independent of the version of the timescale. The set of statements about that concept contained within a particular version of the timescale here are also a versioned graph or description, but the subject of the description (the central 'node') is invariant, and the version information is provided by the context for the set of statements, as described above.

The SKOS and OWL implementations by Ma and colleagues (Ma et al. 2011; Ma et al. 2012) and Perrin et al. (2011) (reviewed by Ma and Fox 2013) provide significant capability. However, these were not based as strictly on the ICS model as the ontology described here, and the SKOS-only implementation does not capture the ordering relationships. Perrin et al. (2011) followed SWEET in using the W3C ontology for time (Hobbs and Pan 2006), but the latter has not progressed beyond 'draft' status. In the ontology described here we chose, rather, to use an OWL implementation of the temporal model from ISO 19108:2002, which provides a more comprehensive set of models not only for temporal geometry, but also for temporal topology and reference systems, all of which are required for a geologic timescale model.

7. Summary and conclusions

We have developed an ontology for the geologic timescale, which combines concepts of temporal topology and hierarchical ordinal reference systems, and links to manifestations in the geologic record that are required to calibrate the timescale, as described in the process of the International Commission for Stratigraphy. The ontology is developed from a UML model previously described by Cox & Richard (2005). It is aligned with an OWL representation of the ISO 19100 harmonized model (Cox 2013), which provides sampling and observation models, as well as basic temporal structures. It is also aligned with SKOS, which supports delivery through generic vocabulary services, and with other common RDF-based vocabularies, including RDFS, FOAF, and Dublin Core.

We have published multiple instances of the timescale, based on versions of the International Chronostratigraphic Chart published over the last 10 years (Cohen et al. 2013). The vocabularies are published through various interfaces: as files for download, at SPARQL endpoints, through a SISSvoc vocabulary interface, and by the URI for each individual era, boundary and supporting concepts used in their definition and calibration. Dereferencing the URI for an era or boundary will return the OWL description of the timescale element corresponding to the most recent version of the International Chronostratigraphic Chart by default, though descriptions from earlier versions are available from parallel services. The intention is that the elements of the vocabulary shall be used to classify geologic datasets, and that the multiple versions will support analysis of the evolution of the geologic timescale.

8. Acknowledgements

Terry Rankine and Jacqui Githaiga (CSIRO), and Stuart WIlliams (Epimorphics) assisted with the SISSvoc interface implementation. Doug Fils (ODP), Marshall Ma (RPI), Jonathan Yu (CSIRO), and Stephan Zednik (RPI), provided useful comments on the use and publication of RDF/OWL. Jim Ogg and Felix Gradstein of ICS provided some advice about the timescale structure. Guillaume Duclaux and Jonathan Yu (CSIRO) helped clarify some aspects of the paper in review. SJDC would like to pay tribute to the late W.B. Harland, who introduced the challenges of formalizing stratigraphy and the timescale in the undergraduate curriculum at Cambridge.

9. References

Beckett D, Berners-Lee T, Prud'hommeaux E, Carothers G (2014) RDF 1.1 Turtle. World Wide Web Consortium

Berrueta D, Phipps J (2008) Best Practice Recipes for Publishing RDF Vocabularies. World Wide Web Consortium

Bizer C, Lehmann J, Kobilarov G, et al. (2009) DBpedia - A crystallization point for the Web of Data. Web Semant 7:154–165. doi: 10.1016/j.websem.2009.07.002

Cohen KM, Finney SC, Gibbard PL, et al. (2013) The ICS International Chronostratigraphic Chart. Episodes 36:199–204.

Cox SJD (2013) An explicit OWL representation of ISO/OGC Observations and Measurements. In: Corcho O, Henson C, Barnaghi P (eds) Proc. 6th Int. Work. Semant. Sens. Networks co-located with 12th Int. Semant. Web Conf. (ISWC 2013). Sun SITE Central Europe, Sydney, Australia, October 22nd, 2013., pp 1–18

Cox SJD, Richard SM (2005) A formal model for the geologic time scale and global stratotype section and point, compatible with geospatial information transfer standards. Geosphere 1:119. doi: 10.1130/GES00022.1

Cox SJD, Richard SM (2014) RDF representation of International Chronostratigraphic Chart (Geologic Timescale). doi: 10.4225/08/537452F354E36

Cox SJD, Yu J, Rankine T (2014) A Linked Data API for SKOS vocabularies. Semant. Web J. submitted:

Fils D, Cervato C, Reed J, et al. (2009) CHRONOS architecture: Experiences with an open-source services-oriented architecture for geoinformatics. Comput Geosci 35:774–782. doi: 10.1016/j.cageo.2008.02.035

Gradstein FM, Ogg JG, Smith AG (2005) Geologic Time Scale 2004. Cambridge University Press

Hobbs JR, Pan F (2006) Time Ontology in OWL. World Wide Web Consortium

ISO/IEC (2005) ISO/IEC 19501:2005 - Information technology -- Open Distributed Processing -- Unified Modeling Language (UML) Version 1.4.2. International Organization for Standardization, Geneva

ISO/TC-211 (2002) ISO 19108:2002 - Geographic information -- Temporal schema. International Organization for Standardization, Geneva

ISO/TC-211 (2011) ISO 19156:2011 - Geographic information -- Observations and measurements. International Organization for Standardization, Geneva

ISO/TC-211 (2014a) ISO/DIS 19150-2:2014 - Geographic information -- Ontology -- Part 2: Rules for developing ontologies in the Web Ontology Language (OWL). International Organization for Standardization, Geneva

ISO/TC-211 (2005a) ISO/TS 19103:2005 - Geographic information -- Conceptual schema language. International Organization for Standardization, Geneva

ISO/TC-211 (2005b) ISO 19109:2005 - Geographic information -- Rules for application schema. International Organization for Standardization, Geneva

ISO/TC-211 (2014b) ISO/DIS 19103 - Geographic information -- Conceptual schema language. International Organization for Standardization, Geneva

ISO/TC-211 (2013) ISO/DIS 19109:2013 - Geographic information -- Rules for application schema. International Organization for Standardization, Geneva

Ma X, Carranza EJM, Wu C, et al. (2011) A SKOS-based multilingual thesaurus of geological time scale for interoperability of online geological maps. Comput Geosci 37:1602–1615. doi: 10.1016/j.cageo.2011.02.011

Ma X, Carranza EJM, Wu C, van der Meer FD (2012) Ontology-aided annotation, visualization, and generalization of geological time-scale information from online geological map services. Comput Geosci 40:107–119. doi: 10.1016/j.cageo.2011.07.018

Ma X, Fox P (2013) Recent progress on geologic time ontologies and considerations for future works. Earth Sci Informatics 6:31–46. doi: 10.1007/s12145-013-0110-x

Miles A, Bechhofer S (2009) SKOS Simple Knowledge Organization System Reference. World Wide Web Consortium

Murphy M, Salvador A (1994) International Stratigraphic Guide - Abridged Edition. International Commission for Stratigraphy and Geological Society of America

Object Management Group (2009) Ontology Definition Metamodel - Version 1.0. Object Management Group, Needham, Mass. USA

Perrin M, Mastella LS, Morel O, Lorenzatti A (2011) Geological time formalization: an improved formal model for describing time successions and their correlation. Earth Sci Informatics 4:81–96. doi: 10.1007/s12145-011-0080-9

Raskin RG, Pan MJ (2005) Knowledge representation in the semantic web for Earth and environmental terminology (SWEET). Comput Geosci 31:1119–1125. doi: 10.1016/j.cageo.2004.12.004

Remane J, Bassett MG, Cowie JW, et al. (1996) Revised guidelines for the establishment of global chronostratigraphic standards by the International Commission on Stratigraphy (ICS). Episodes 19:77–81.

Richard SM, CGI-Interoperability-Working-Group (2007) GeoSciML – A GML Application for Geoscience Information Interchange. In: Soller D (ed) Digit. Mapp. Tech. '06 — Work. Proceedings, USGS Open File Rep. 2007-1285. United States Geological Survey, Columbus, OH., pp 47–59

Shergold JH, Cooper RA (2005) The Cambrian Period. A Geol. Time Scale 2004. Cambridge University Press, Cambridge, UK, pp 147–164

W3C OWL Working Group (2012) OWL 2 Web Ontology Language Document Overview (Second Edition). World Wide Web Consortium