isa

Interoperability Solutions
for European Public Administrations

CORE VOCABULARIES SPECIFICATION

CORE BUSINESS VOCABULARY

CORE PERSON VOCABULARY

CORE LOCATION VOCABULARY

JOINING UP GOVERNMENTS

DOCUMENT METADATA

Property	Value
Release date	*11/05/2012*
Status:	*Final*
Version:	*1.00*

Statement of copyright:

This specification was prepared for the ISA programme by:
PwC EU Services

Disclaimer:

TABLE OF CONTENTS

1. INTRODUCTION

1.1 ABOUT THE ISA PROGRAMME

This specification has been created as part of Action 1.1 [A1.1] of the Interoperability solutions for European public administrations (ISA) programme of the European Commission (EC). This programme funds initiatives to foster the efficient and effective cross-border electronic interactions between European public administrations. Action 1.1 of this programme is targeted towards improving the semantic interoperability of European e-Government systems. It addresses these by encouraging the sharing and reuse of semantic assets. As part of Action 1.1, the ISA Programme intends to build consensus on a number of e-Government Core Vocabularies for public sector information exchange.

1.2 MOTIVATION

The **metadata** and **reference data** used in electronic public services across Europe most often has a very specific context. Attaining consensus on common metadata and reference data for these electronic services is a *critical* step towards semantic interoperability. Unfortunately, consensus building is hindered by the diverse cultural, multi-lingual, legal, and organisational contexts of these e-Government services. To alleviate this problem, consensus building should start at a higher level of abstraction that surpasses the contexts of individual electronic public services, and thus the cultural, lingual, legal, and organisational differences of individual countries. In particular, consensus can be more easily attained on the semantics of a *small* set of fundamental concepts, for which less divergent opinions exist [EGOV-CV]. These concepts are what we describe as Core Vocabularies.

1.3 TERMINOLOGY

This document uses the following terminology:

Semantic interoperability is defined as the ability of information and communication technology (ICT) systems and the business processes they support to exchange data and to enable the sharing of information and knowledge: *Semantic Interoperability enables systems to combine received information with other information resources and to process it in a meaningful manner* (European Interoperability Framework 2.0[1]). It aims at the mental representations that human beings have of the meaning of any given data.

A **Core Vocabulary** is a simplified, reusable, and extensible data model that captures the fundamental characteristics of an entity in a context-neutral fashion [EGOV-CV]. Well known examples of existing Core Vocabularies include the Dublin Core Metadata Set [DC]. Such Core Vocabularies are the starting point for agreeing on new semantic interoperability assets and defining mappings between existing assets. Semantic interoperability assets that map to or

[1] http://ec.europa.eu/isa/documents/isa_annex_ii_eif_en.pdf

extend such Core Vocabularies are the *minimum required* to guarantee a level of cross-domain and cross-border interoperability that can be attained by public administrations.

1.4 THIS SPECIFICATION

The Core Vocabularies defined here present concepts and terms that can be used to describe three things that are closely related:

- People
- Locations
- Businesses

This is obviously a very short list, one that excludes many other domains that are no less interconnected such as, accounts and payments, projects and processes, buildings, education and healthcare organisations. Any of these and many more could have been included and were considered. However, it was decided to focus, initially at least, on these three areas as there was evidence from some public sector data repositories and broader discussions that these were the highest priority areas. Furthermore, the links between them are particularly strong so that the three create a single coherent set.

In developing these three core vocabularies the working group consciously limited its work further. The vocabulary for describing a person relates to the **natural person**, i.e. the individual as opposed to any role they may play in society or the relationships they have to other people, organisations and property; all of which contribute significantly to the broader concept of *identity*. In describing businesses, the working group focussed solely on **legal entities**, that is, trading bodies that are formally registered with the relevant authority and that appear in business registers. This excludes sole traders, virtual organisations and other types of 'agent' that are able to do business. The broadest vocabulary is that for location which covers **places**, **addresses**, and geographical **geometries**. Specific vocabulary terms that were considered but not included are described in Section 6.

1.5 RELATED WORK

Care has been taken to ensure that the defined terms are most likely to match those used by others, thus maximising the interoperability and re-usability of data. A minimum number of new terms has been defined with the majority of those called for in the conceptual model being re-used from existing, widely used, vocabularies. The existing vocabularies referred to include:

- the UN Centre for Trade Facilitation and e-Business [UN/CEFACT]
- Friend of a Friend [FOAF]
- INSPIRE Addresses Technical Guidelines [INADR]
- W3C Point of Interest Core [POI]
- xEBR Core reference Taxonomy created by XBRL Europe [xEBR]
- schema.org

- The work of ISO/TC 211, Geographic information/Geomatics, and its standards, such as ISO/TS 19104:2008, Geographic information – Terminology, ISO 19112:2003, Geographic information -- Spatial referencing by geographic identifiers, ISO 19152, Geographic information – Land administration data model, and ISO 19160, Addressing.
- OGC's GML
- OASIS CIQ xPRL V3.0 (extensible Party Relationships Language) and xAL.

Further information on the approach taken is provided in Section 9 Background and Objectives.

2. CONFORMANCE STATEMENT

Conformance with this specification can be achieved in a number of ways as detailed below. For specific notes on conformance of address data with the INSPIRE Directive, see section 3.5.1.

2.1 DATA PUBLISHING

Conceptual conformance, meaning that data is published along with a human readable mapping from the terms used by the publisher to the terms defined in this document.

Conceptual conformance+ meaning that data is published along with a machine processable mapping, such as an XSLT, from the terms used by the publisher to the terms defined by this document.

RDF conformance means that data is published using the RDF classes and properties specified this document and its associated RDF namespace document.

XML conformance means that data is published using the XML elements and types specified in this document and its associated XML namespace document.

Full data publishing conformance means that data is published in both RDF and XML formats with the relevant format returned from the server through content negotiation.

2.2 DATA CONSUMING

A consumer of data published using the Core Vocabularies vocabulary must understand all the terms defined below in one of three ways which should be declared when claiming conformance:

RDF conformance, meaning that RDF data published using any term in the vocabulary will be consumed and processed accordingly;

XML conformance, meaning that XML data published using any term in the vocabulary will be consumed and processed accordingly;

XML and RDF conformance, meaning that both RDF and XML data published using any term in the vocabulary will be consumed and processed accordingly.

2.3 PRIVACY

Access control and provenance is out of scope for the current work but data publishers and consumers should not take use of any core vocabulary as an automatic 'right to share.' Their use has no effect on ownership and access rights. Personal data is, by its nature, sensitive and must be handed accordingly. Data provenance and security is not covered by this specification but it is recognised as being of very high importance.

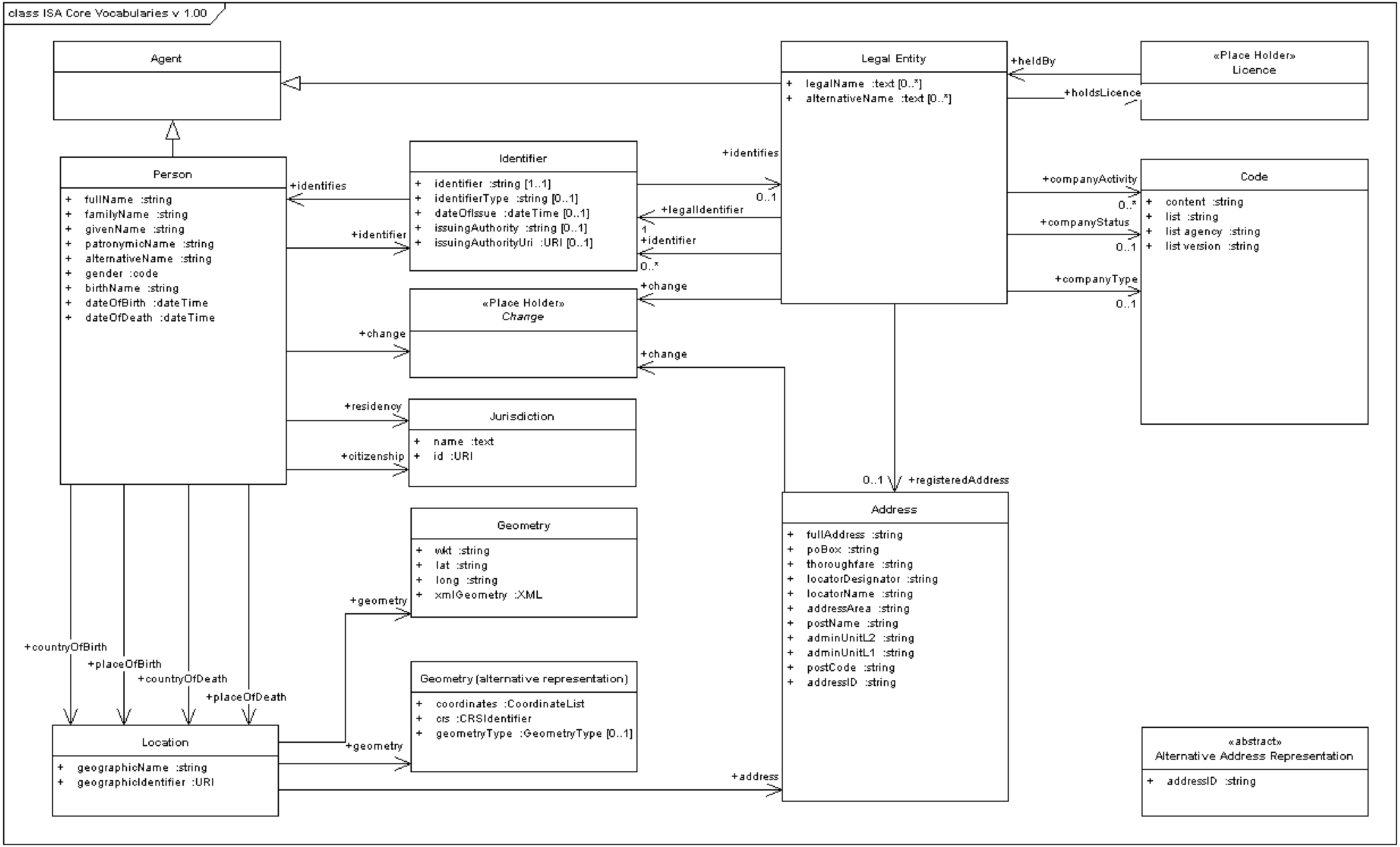

Figure 1 UML diagram for the ISA Core Vocabularies

3. CONCEPTUAL MODEL

The conceptual model presented in Figure 1 is independent of any technology that may be used to represent it. It describes the minimal set of classes, relationships and properties necessary to describe a natural person, a business that has legal entity status, and a location. Cardinality restrictions are only applied to the Business Core Vocabulary where the use case demands it, otherwise no such restrictions are applied to ensure maximum flexibility.
In the following sections, each class is defined, followed by definitions of its properties (data type properties) and relationships (object properties).

3.1 PERSON

The Person Class is a sub class of the more general 'Agent' class that encompasses organisations, legal entities, groups etc. - any entity that is able to carry out actions.

For avoidance of doubt, the word 'relationships' is used throughout this document as a synonym for an object property. This vocabulary does not cover the societal relationships that an individual person has with, for example, other members of their family or their employer. The concern here is with the key aspects that define the natural person. The dividing line between describing a natural person and describing their identity is imprecise and some properties that the working group felt were 'just' out of scope are among those discussed in section 6.

The data type properties of the Person class do not have any cardinality restrictions and as such all are optional. However, guidance is provided for the usage of each property in the following sections.

3.1.1 full name

Property	Abstract Data Type
fullName	String

The full name (fullName) property contains the complete name of a person as one string. In addition to the content of given name, family name and, in some systems, patronymic name, this can carry additional parts of a person’s name such as titles, middle names or suffixes like “the third” or names which are neither a given nor a family name.

The Full Name is the most reliable label for an individual and as such its use is strongly encouraged, irrespective of whether that name is broken down using the more granular elements. It is anticipated that some systems will *only* provide or process the full name of a person. Where an individual has more than one full legal name (a relatively rare but not unknown phenomenon), the full name property can be used more than once. In this case, however, the granular name elements should not be used since the intention is that these provide a breakdown of the full name and it will not be clear of which full name this is true.

Note that the vocabulary provides an alternative name property (3.1.5). This allows name(s) to be recorded that have no legal status but that nevertheless are the names by which an individual is generally known.

Reasoning

A name usually sticks with a person for a long time period. In some European countries a name may only be changed according to certain laws and life events, e.g. marriage. The name denominates a natural person even if he/she changes their address. Documents like birth certificate or diploma usually don't carry an address but always the name. Thus the name is one of the core attributes. However it is not sufficient to identify a person since there are combinations of very common names like Smith in the UK, Meier in Germany, or Li in China.

3.1.2 given name

Property	Abstract Data Type
givenName	String

A given name, or multiple given names, are the denominator(s) that identify an individual within a family. These are given to a person by his or her parents at birth or may be legally recognised as 'given names' through a formal process. All given names are ordered in one field so that, for example, the Given Name for Johan Sebastian Bach is 'Johan Sebastian.'

3.1.3 family name

Property	Abstract Data Type
familyName	String

A family name is usually shared by members of a family. This attribute also carries prefixes or suffixes which are part of the Family Name, e.g. “de Boer”, “van de Putte”, “von und zu Orlow”. Multiple family names, such as are commonly found in Hispanic countries, are recorded in the single Family Name field so that, for example, Miguel de Cervantes Saavedra's Family Name would be recorded as "de Cervantes Saavedra.

3.1.4 patronymic name

Property	Abstract Data Type
patronymicName	String

Patronymic names are important in some countries. Iceland does not have a concept of 'family name' in the way that many other European countries do, for example. Erik Magnusson and Erika Magnusdottir are siblings, both offspring of Mangnus, irrespective of his patronymic name. In Bulgaria and Russia, patronymic names are in every day usage, for example, the Sergeyevich in 'Mikhail Sergeyevich Gorbachev.'

Note that patronymic names refer to a father's given name, not the family name inherited from the mother and father as is the case in countries such as Spain and Portugal. Again referring to the example of Miguel de Cervantes Saavedra's, the patronymic name element would be unused.

3.1.5 alternative name

Property	Abstract Data Type
alternativeName	String

Any name by which an individual is known other than their full name. Many individuals use a short form of their name, a 'middle' name as a 'first' name or a professional name. For example, the British politician and former UN High Representative for Bosnia and Herzegovina, Jeremy John Durham Ashdown, Baron Ashdown of Norton-sub-Hamdon, is usually referred to simply as 'Paddy Ashdown' or 'Lord Ashdown.' See Appendix B which uses Paddy Ashdown as an example of how many of the properties of the Person class, including alternative name, should be used.

Note

It is not the role of the alternative name property to record nick names, pet names or other 'familiar names' that will be of no consequence in public sector data exchange. Furthermore, some individuals have more than one legal name in which case the full name property should be used multiple times. Alternative name gives a means of recording names by which an individual is generally known, or professionally known, even though such names are no more than secondary from a legal point of view.

3.1.6 gender

Relationship	Abstract Data Type
gender	Code

The gender of an individual should be recorded using a controlled vocabulary that is appropriate for the specific context. In some cases the chromosomal or physical state of an individual will be more important than the gender that they express, in others the reverse will be true. What is always important is that the controlled vocabulary used to describe an individual's gender is stated explicitly. To do this we use the Code datatype which is described in detail in section 5.3.

Four examples of controlled vocabularies that can be used to describe a person's gender are detailed below (there are others too). As is clear, these vocabularies share many features and can generally be mapped from one to another fairly easily.

ISO/IEC 5218:2004 [ISO5218]

0 not known;

1 male;

2	female;
9	not applicable.

Eurostat SCL - Sex [SCLS]

F	female
M	male
OTH	other
UNK	unknown
NAP	not applicable

HL7

F	Female
M	Male
UN	Undifferentiated (the gender of a person could not be uniquely defined as male or female, such as hermaphrodite).

OID: 2.16.840.1.113883.5.1
URI: http://www.oid-info.com/get/2.16.840.1.113883.5.1

SDMX

F	Female
M	Male
U	Not specified or unknown
N	Not applicable
T	Total

See SDMX Cross-Domain Code Lists [SDMX CDL]

SDMX provides a full XML schema for its code lists (see http://sdmx.org/?page_id=16) as well as an RDF schema that provides URIs of the form http://purl.org/linked-data/sdmx/2009/code#sex-{x} where {x} represents the character from the list above, e.g. http://purl.org/linked-data/sdmx/2009/code#sex-F.

3.1.7 birth name

Property	Abstract Data Type
birthName	String

All data associated with an individual are subject to change. Names can change for a variety of reasons, either formally or informally, and new information may come to light that means that a correction or clarification can be made to an existing record. Birth names tend to be persistent however and for this reason they are recorded by some public sector information systems. There is no granularity for birth name - the full name should be recorded in a single field.

3.1.8 date of birth, date of death

Property	Abstract Data Type
dateOfBirth dateOfDeath	dateTime

A date that specifies the birth date (death date) of a person. See section 5 for details of the dateTime data type.

3.1.9 country of birth, country of death

Relationship	Abstract Data Type
countryOfBirth countryOfDeath	Location

A person's Country of Birth and Death is defined using the Location class which is associated via the appropriate relationship.

The Location Class, defined in Section 3.4, has two properties: a Geographic Name and a Geographic Identifier. Plain codes like "DE" should be provided as values for Geographical Names whereas URIs should be provided as value of the Geographical Identifier. Ideally, provide both.

Providing a simple country name is problematic and should be avoided whereas using a standardised system that allows the use of a code list for country names has a lot of potential for increasing semantic interoperability.

Known diversity that one has to deal with when exchanging country names between different communication partners without relying on an agreed code list are:

a. Long form vs. short form of a country name (e.g. “Federal Republic of Germany” vs. Germany),
b. Different languages (Italy vs. Italia),
c. Historic name vs. Current name (Burma vs. Myanmar),
d. Ambiguity of similar sounding countries (“Republic of the Congo” vs. “Democratic Republic of the Congo”)

The Publications Office of the European Union recommends and uses ISO 3166-1 codes for countries in all cases except two:

- use 'UK' in preference to the ISO 3166 code GB for the United Kingdom;
- use 'EL' in preference to the ISO 3166 code GR for Greece.

See http://publications.europa.eu/code/en/en-5000500.htm for details of the OPOCE's full list of countries, codes, currencies and more.

Where a country has changed its name or no longer exists (such as Czechoslovakia, Yugoslavia etc.) use the ISO 3166-3 code [ISO 3166-3]

3.1.10 place of birth, place of death

Relationship	Target Class
placeOfBirth placeOfDeath	Location

The Place of Birth and Place of Death are given using the Location class which is associated via the appropriate relationship. As detailed in Section 3.4, the Location Class has two properties:

- The geographic name of the place, which is given as a string such as "Amsterdam" or "Valetta";
- an identifier, such as a geonames URI http://sws.geonames.org/2759794 (which identifies Amsterdam) or http://sws.geonames.org/2562305 (which identifies Valetta).

The use of identifiers is preferred as these are unambiguous, however, public sector data typically uses simple names to record places and this is fully supported.

3.1.11 citizenship, residency and Jurisdiction

Relationship	Abstract Data Type
citizenship residency	Jurisdiction

The citizenship relationship links a Person to a Jurisdiction that has conferred citizenship rights on the individual such as the right to vote, to receive certain protection from the community or the issuance of a passport. Multiple citizenships are recorded as multiple instances of the citizenship relationship. Residency typically provides an individual with a subset of the rights of a citizen. Both citizenship and residency link a Person with a Jurisdiction.

The Jurisdiction class has just two properties:

Property	Abstract Data Type
name	Text
id	URI

The name is simply a string that identifies the jurisdiction, typically a country, with or without a language tag. The value for the id property is a URI for that jurisdiction.

Reasoning

A person has one, multiple or even no citizenship or residency status. Citizenship is information needed by many cross-border use cases and is a legal status as opposed to the more culturally-focussed and less well-defined term "nationality." As with Country of Birth and Death, using plain text names of countries is strongly discouraged.

3.1.12 identifier

Relationship	Abstract Data Type
identifier	Identifier

For many systems, the identifier is the key piece of information about an individual and therefore an important part of the core person vocabulary. However, all identifiers are context-specific and when exchanging data between systems it is important to provide additional information that makes this explicit. An individual may have any number of identifiers. For this reason the Identifier class is used (defined in Section 3.3).

The Identifier class allows an identifier to be given as well as information about which authority issued it and, if needed, when.

3.2 LEGAL ENTITY

This is the key class for the Business Core Vocabulary and represents a business that is legally registered. In many countries there is a single registry although in others, such as Spain and Germany, multiple registries exist. A Legal Entity is able to trade, is legally liable for its actions, accounts, tax affairs etc.

This makes legal entities distinct from the concept of organisations, groups or sole traders. Many organisations exist that are not legal entities yet to the outside world they have staff, hierarchies, locations etc. Other organisations exist that are an umbrella for several legal entities (universities are often good examples of this). This vocabulary is concerned solely with registered legal entities and does not attempt to cover all possible trading bodies.

Legal Entity is a sub class of the more general 'Agent' class that does encompass organisations, natural persons, groups etc. - i.e. an Agent is any entity that is able to carry out actions. Unlike the properties of the Person class or the location classes, properties of the Legal Entity and Identifier classes *do* have cardinality constraints although these are as unrestrictive as possible.

3.2.1 legal name

Property	Abstract Data Type	Cardinality
legalName	Text	[0..*]

The legal name of the business. A business might have more than one legal name, particularly in countries with more than one official language. In such cases, and where the encoding technology allows, the language of the string should be identified (see Section 5.2).

3.2.2 alternative name

Attribute	Abstract Data Type	Cardinality
alternativeName	Text	[0..*]

Some jurisdictions recognise concepts such as a trading name or alternative forms of a legal entity's name. The Alternative Name property can be used to record such names but should not be used to record translations of the primary legal name. Where more than one legal name exists and where they have equal standing but are expressed in different languages, identify the language used in each of the multiple legal names (see previous section).

It is notable that some jurisdictions regard the use of any name other than the primary Legal Name as suspicious.

3.2.3 company type

Relationship	Abstract Data Type	Cardinality
companyType	Code	[0..1]

This property records the type of company. Familiar types are SA, PLC, LLC, GmbH etc. At the time of publication, there is no agreed set of company types that crosses borders. The term 'SA' is used in Poland and France for example although they mean slightly different things. The UK's LLP and Greece's EPE provide further example of close, but not exact, matches.

That said, each jurisdiction will have a limited set of recognised company types and these should be used in a consistent manner. It is to be hoped that further work can be carried out in this area in the near future. For now, the company type and the vocabulary from which the term comes should be recorded. See section 5.3 for guidance on using controlled vocabularies.

3.2.4 company status

Relationship	Abstract Data Type	Cardinality
companyStatus	Code	[0..1]

Recording the status of a company presents the same issues as its type. The terms 'insolvent', 'bankrupt' and 'in receivership,' for example, are likely to mean slightly different things with different legal implications in different jurisdictions. Again, at the time of publication, there is no agreed vocabulary for recording the status of a company although XBRL Europe [xEBR] is in the process of defining such a vocabulary.

Taking the existing XBRL work as a starting point, however, the term 'Normal Activity' does appear to have cross-border usefulness and this should be used in preference to terms like 'trading' or 'operating.'

Best Practice for recording various other status levels is to use the relevant jurisdiction's terms and to identify the controlled vocabulary used. See section 5.3 for guidance on using controlled vocabularies.

3.2.5 company activity

Relationship	Abstract Data Type	Cardinality
companyActivity	Code	[0..*]

The activity of a company should be recorded using a controlled vocabulary. Several such vocabularies exist, many of which map to the UN's ISIC codes [ISIC4]. Where a particular controlled vocabulary is in use within a given context, such as SIC codes in the UK, it is acceptable to use these, however, the preferred choice for European interoperability is NACE [NACE]. Again, section 5.3 offers guidance on using controlled vocabularies.

3.2.6 legal identifier

Relationship	Target Class	Cardinality
legalIdentifier	Identifier	[1..1]

The legal status of a business is conferred on it by an authority within a given jurisdiction. The Legal Identifier is therefore a fundamental relationship between a legal entity and the authority with which it is registered. The details of the registration are provided as properties of the Identifier class. The Core vocabulary sets no restriction on the type of legal identifier. In many countries, the business register's identifier is the relevant data point. The tax number often fulfils this function in Spain.

The cardinality of this relationship is 1..1, i.e. a legal entity shall have a unique legal identifier and a legal identifier shall identify only one legal entity.

In RDF terms, legalIdentifier is a sub property of identifier (see next section) with specific semantics.

Rationale

Core Vocabularies do not normally demand that any specific field (property/relationship) is defined. However, a legal entity can only have that status if it is given by a suitable authority. Therefore the working group resolved that a Legal Entity Class could not sensibly be defined without the legal identifier.

3.2.7 identifier

Relationship	Target Class	Cardinality
identifier	Identifier	[0..*]

The identity relation links a resource to any formally issued identifier for that resource other than the one that confers legal status upon it.

Legal Entities, people, organisations and other Agents may have any number of identifiers (but only one legal identifier). For example, in many jurisdictions, a business will have one or more tax numbers associated with them which do not, by themselves, confer legal entity status. An individual may be issued with identifiers for everything from social security to club membership. The identifier relationship must not be used to link to the identifier issued by the authority that conferred legal entity status on a business.

In RDF terms, the identifier property is not constrained by its domain so that any resource can be linked to an Identifier.

3.2.8 registered address

Relationship	Target Class	Cardinality
registeredAddress	Address	[0..1]

In almost all jurisdictions, legal entities must register a public address. This may or may not be the actual address at which the legal entity does its business, it is commonly the address of their lawyer or accountant, but it is the address to which formal communications can be sent.

In RDF terms, registeredAddress is a sub property of the more general address property that links any resource to an associated address without further semantics. See Section 3.4.3.

3.2.9 legal entity

Relationship	Target Class	Cardinality
legalEntity	Legal Entity	[0..*]

The legal entity relationship can be used to link any resource to a Legal Entity Class. This is useful, for example, where an organisation includes one or more legal entities. The Dublin Core term isPartOf is a suitable inverse of this relationship [DC].

3.3 IDENTIFIER

The Identifier class represents any identifier issued by any authority, whether a government agency or not. It captures the identifier itself, the type of identifier, and details of the issuing authority, the date on which the identifier was issued. The Identifier class is based on the

UN/CEFACT class of the same name [UN/CEFACT] and is defined under the ADMS namespace [ADMS]. See section 5.4 which provides full details of the Identifier class as a complex data type.

3.3.1 identifies

Relationship	Target Class	Cardinality
identifies	Resource	[0..1]

The identifies relationship links an Identifier class to the resource it identifies. In the context of the Business Core Vocabulary this will be a Legal Entity but the class can be used to identify any resource, including any sub class of Agent.

3.4 LOCATION

Locations can be described in three principal ways: by using a place name, a geometry or an address. The specific context will determine which method of describing a location is most appropriate. The Location Core Vocabulary provides structure for all three.

ISO 19112 defines a location as "an identifiable geographic place." With this in mind, "Eiffel Tower", "Madrid" and "California" are all locations and this is a common way of representing locations in public sector data, i.e. simply by using a recognised name. Such identifiers are common although they can be highly ambiguous as many places share the same or similar names.

In addition to a simple (string) label or name for a Location, this vocabulary defines a property that allows a Location to be defined by a URI, such as a GeoNames or DBpedia URI.

No cardinality constraints are placed on any property of the Location, Address or Geometry classes so as to maximise flexibility. A single address may be defined in different ways, a geometry may be defined using different coordinate reference systems and a single place may have no recognised name or multiple names. The Location Core vocabulary makes a minimum number of assumptions about what data will be encoded. However, it is clearly nonsense to define any of the location classes without any properties or to provide multiple instances of the same property with conflicting values.

3.4.1 geographic name

Attribute	Abstract Data Type
geographicName	String

The INSPIRE Data Specification on Geographical Names [INGN] provides a detailed model for describing a 'named place', including methods for providing multiple names in multiple scripts.

This is beyond what is necessary for the Core Location vocabulary but, importantly, the concept of a geographic name used here is consistent.

A geographic name is a proper noun applied to a spatial object. Taking the example used in the INSPIRE document (page 18), the following are all valid geographic names for the Greek capital:

- Αθnva (the Greek endonym written in the Greek script)
- Athína (the standard Romanisation of the endonym)
- Athens (the English language exonym)

INSPIRE has a detailed (XML-based) method of providing metadata about a geographic name and in XML-data sets that may be the most appropriate method to follow. When using the Location Core vocabulary in data sets that are not focussed on environmental/geographical data (the use case for INSPIRE), the Code datatype (section 5.3) or a simple language identifier (section 5.2) may be used to provide such metadata.

The country codes defined in ISO 3166 may be used as geographic names and these are generally preferred over either the long form or short form of a country's name (as they are less error prone). As noted in Section 3.1.9, the Publications Office of the European Union recommends the use of ISO 3166-1 codes for countries in all cases **except** two:

- use 'UK' in preference to the ISO 3166 code GB for the United Kingdom;
- use 'EL' in preference to the ISO 3166 code GR for Greece.

Where a country has changed its name or no longer exists (such as Czechoslovakia, Yugoslavia etc.) use the ISO 3166-3 code [ISO 3166-3]

3.4.2 geographic identifier

Attribute	Abstract Data Type
geographicIdentifier	URI

A URI that identifies the location.

GeoNames.org provides stable, widely recognised identifiers for more than 10 million geographical names that can be used as links to further information. For example, http://sws.geonames.org/593116/ identifies the Lithuanian capital Vilnius. Unfortunately these URIs cannot easily be automatically deduced since the URI scheme uses simple numeric codes. Finding a GeoNames identifier for a Location is almost always a manual process. Where such identifiers are known or can be found, however, it is recommended that they be used.

Where the Location Class is used to identify a country, if the geonames URI is not known, the recommendation is to use DBpedia URIs of the form http://dbpedia.org/resource/ISO_3166-1:XX where XX is the ISO 3166 two character code for the country [DBpedia].

As noted in the previous section and in 3.1.9, the EU's Publication Office diverges from ISO 3166-1 and uses EL and UK for Greece and the United Kingdom respectively. DBpedia sticks to the ISO codes and so the correct URIs for these countries are:

- http://dbpedia.org/resource/ISO_3166-1:GR
- http://dbpedia.org/resource/ISO_3166-1:GB

even when the geographic name is given as EL or UK.
The use of a URIs has added advantages:

a. it can be used by automated systems to look up additional data (linked data);
b. a triple store may store only one copy of the URI, whereas if a string is used, a copy of that string is always stored for each and every person in the database. Thus, in large data sets, the saving on memory capacity and the improvement in transmission efficiency can be substantial.

3.4.3 address

Relationship	Target Class
address	Address

The address relationship associates any resource with the Address Class (i.e. anything can be linked to its address using this property). Asserting the address relationship implies that the resource has an address. Although not shown explicitly in Figure 1, this means that address may be used to link a Person to their address.

3.4.4 geometry

Relationship	Target Class
geometry	Geometry

The geometry property may have as its range either a literal (such as WKT, GML or KML) or a geometry class (e.g., ogc:Geometry and its subclasses, geo:Point, schema:GeoCoordinates and schema:GeoShape, a GeoHash URI reference). See also the Geometry class (section 3.6).

3.4.5 location

Relationship	Target Class
location	Location

The location relationship associates any resource with the Location Class. Asserting the location relationship implies only that the domain has some connection to a Location in time or space. It does not imply that the resource is necessarily at that location at the time when the assertion is made. It is a super property of country of birth/death and place of birth/death (sections 3.1.9, 3.1.10).

3.5 ADDRESS

The representation of addresses varies widely from one country's postal system to another. Even within countries, there are almost always examples of addresses that do not conform to the stated national standard. At the time of publication, work is progressing on ISO 19160-1 that defines a method through which different addresses can be converted from one conceptual model to another.

For this specification we have drawn heavily on the INSPIRE Address Representation data type [INADR]. The relevant properties, all of which can be found in the INSPIRE Address Representation Class, are detailed in the following sections but in brief they are:

- PO Box (a specialisation of locator designator)
- Thoroughfare (a road, a waterway etc.)
- Locator designator (a building number, entrance number etc.)
- Locator name (a proper name for a building or room within a building)
- Address area (usually a city area or village)
- Locality (usually a town)
- Admin unit level 2 (usually a county or state)
- Admin unit level 1 (almost always a country)
- Post Code

It is noteworthy that **if** an address is provided using the **detailed** breakdown suggested by these properties for this class, then it will be **INSPIRE-conformant**. To this very granular set of properties we add two further properties:

- full address (the complete address as a formatted string)
- addressID (a unique identifier for the address)

The first of these allows publishers to simply provide the complete address as one string, with or without formatting. This is analogous to vCard's label property.

The addressID is part of the INSPIRE guidelines and provides a hook that can be used to link the address to an alternative representation, such as vCard (see Section 7.1) or OASIS xAL [XAL].

3.5.1 INSPIRE Conformance

As already noted, the granular address properties (PO Box through to Post Code) are taken from the INSPIRE address guidelines, as is address ID. An address that is provided using these properties will be INSPIRE conformant. The additional property of full address is not part of the INSPIRE Address guidelines.

In order to maximise flexibility in this Core Vocabulary, there are no cardinality constraints on any location property. Therefore, from a technical point of view, a data set that declares an instance of an Address class with no properties at all would be conformant to the Location Core Vocabulary. Such a data set would, however, be meaningless. Publishers should therefore decide whether or not to publish data that is conformant with INSPIRE.

To achieve INSPIRE conformance using the Core Location vocabulary, the detailed address information must be given, notably including separate values for the thoroughfare and the locator designator or locator name. The full address may also be given but is likely to be ignored by INSPIRE-based systems.

If less detailed data is available then any of the properties may be used, including most notably, the full address property which takes a single string representation of an address. If address data is available in a different format, then use the address ID as a link between the two. An Address that is given in terms of values for just these extension properties is conformant with the Core Location vocabulary but is not conformant with INSPIRE.

3.5.2 full address

Attribute	Abstract Data Type
fullAddress	String

The complete address written as a string, with or without formatting. Use of this property is recommended as it will not suffer any misunderstandings that might arise through the breaking up of an address into its component parts.

This property is analogous to vCard's label property but with two important differences:

1. formatting is not assumed so that, unlike vCard label, it may not be suitable to print this on an address label.
2. vCard's label property has a domain of vCard Address. The fullAddress property has no such restriction.

3.5.3 po box

Attribute	Abstract Data Type
poBox	String

The Post Office Box number. INSPIRE's name for this is "postalDeliveryIdentifier" for which it uses the locator designator property with a type attribute of that name. This vocabulary separates out the Post Office Box for greater independence of technology.

3.5.4 thoroughfare

Attribute	Abstract Data Type
thoroughfare	String

Quoting from the INSPIRE guidelines, a thoroughfare is: "an address component that represents the name of a passage or way through from one location to another. A thoroughfare is not necessarily a road, it might be a waterway or some other feature.

3.5.5 locator designator

Attribute	Abstract Data Type
locatorDesignator	String

The locator designator is defined by the INSPIRE guidelines as "a number or a sequence of characters that uniquely identifies the locator within the relevant scope(s). The full identification of the locator could include one or more locator designators." In simpler terms, this is the building number, apartment number, etc. For an address such as "Flat 3, 17 Bridge Street", the locator is "flat 3, 17."

3.5.6 locator name

Attribute	Abstract Data Type
locatorName	String

Again quoting the INSPIRE guidelines, locator name is defined as: "Proper noun(s) applied to the real world entity identified by the locator. The locator name could be the name of the property or complex, of the building or part of the building, or it could be the name of a room inside a building.

The key difference between a locator and a locator name is that the latter is a proper name and is unlikely to include digits. For example, "Shumann, Berlaymont" is a meeting room within the European Commission headquarters for which locator name is more appropriate than locator.

3.5.7 address area

Attribute	Abstract Data Type
addressArea	String

Taking the definition from INSPIRE, the address area is: the name or names of a geographic area or locality that groups a number of addressable objects for addressing purposes, without being an administrative unit." This would typically be part of a city, a neighbourhood or village.

3.5.8 post name (city)

Attribute	Abstract Data Type
postName	String

The key postal division of the address, usually the city. (INSPIRE's definition is "One or more names created and maintained for postal purposes to identify a subdivision of addresses and postal delivery points.)

3.5.9 admin unit level 2 (county/region/state)

Attribute	Abstract Data Type
adminUnitL2	String

The region of the address, usually a county, state or other such area that typically encompasses several localities.

3.5.10 admin unit level 1 (country)

Attribute	Abstract Data Type
adminUnitL1	String

The uppermost administrative unit for the address, almost always a country.

Best practice is to use the ISO 3166-1 code but if this is inappropriate for the context, country names should be provided in a consistent manner to reduce ambiguity. For example, either write 'United Kingdom' or 'UK' consistently throughout the data set and avoid mixing the two.

3.5.11 post code

Attribute	Abstract Data Type
postCode	String

The post code (a.k.a postal code, zip code etc.). Post codes are common elements in many countries' postal address systems.

3.5.12 address id

Attribute	Abstract Data Type
addressId	String

The concept of adding a globally unique identifier for each instance of an address is a crucial part of the INSPIRE data spec. A number of EU countries have already implemented an ID (a

UUID) in their Address Register/gazetteer, among them Denmark. OASIS xAL also includes an address identifier.

It is the address Identifier that allows an address to be represented in a format other than INSPIRE whilst remaining conformant to the core vocabulary. See Section 7.1.

3.6 GEOMETRY

The Geometry Class denotes the notion of geometry at a conceptual level, and can be encoded in different formats including WKT, GML, KML, RDF+WKT/GML (GeoSPARQL), RDF (WGS84 lat/long, schema.org) and GeoHash URI references.

Conceptually, the class has the following properties:

- coordinates (which gives the coordinate list);
- crs (an identifier for the coordinate reference system) ;
- the geometry type (point, line or polygon).

These are shown in Figure 1 in the alternative representation of the class. The conceptual model of the Geometry class itself shows a number of properties too (wkt, lat/long etc.). We do not define any of these properties in the RDF or XML schemas as, in practice, geometry encodings all include these properties and the locn:geometry property is sufficient (section 3.4.4) .

3.6.1 Geometry RDF Examples

The following examples (written in RDF Turtle) try to make this clear:

WKT (GeoSPARQL)

```
<ex:a> a locn:Location ;
 locn:geometry
"<http://www.opengis.net/def/crs/OGC/1.3/CRS84> Point(-0.001475
51.477811)"^^ogc:WKTLiteral .
```

GML

```
<ex:a> a locn:Location ;
 locn:geometry
"<gml:Point srsName="http://www.opengis.net/def/crs/OGC/1.3/CRS84">
  <gml:coordinates>-0.001475, 51.477811
  </gml:coordinates></gml:Point>"^^ogc:GMLLiteral .
```

RDF+WKT (GeoSPARQL)

```
<ex:a> a locn:Location ;
 locn:geometry [ a ogc:Point; ogc:asWKT
  "<http://www.opengis.net/def/crs/OGC/1.3/CRS84>
  Point(-0.001475 51.477811)"^^ogc:WKTLiteral ] .
```

RDF+GML (GeoSPARQL)

```
<ex:a> a locn:Location ;
 locn:geometry
 [ a ogc:Point;
  ogc:asGML "<gml:Point
  srsName="http://www.opengis.net/def/crs/OGC/1.3/CRS84">
  <gml:coordinates>-0.001475, 51.477811
  </gml:coordinates></gml:Point>"^^ogc:GMLLiteral ] .
```

RDF (WGS84 lat/long)

```
<ex:a> a locn:Location ;
 locn:geometry [ a geo:Point;
                geo:lat "51.477811"; geo:long "-0.001475" ] .
```

RDF (schema.org)

```
<ex:a> a locn:Location ;
 locn:geometry [ a schema:GeoCoordinates; schema:latitude "51.477811";
    schema:longitude "-0.001475" ] .
```

URI reference (GeoHash)

```
<ex:a> a locn:Location ;
 locn:geometry <http://geohash.org/gcpuzgnzvxkp> .
```

3.6.2 Geometry XML Examples

WKT (GeoSPARQL)

```
< Geometry>
      <GeometryExpression>
            <http://www.opengis.net/def/crs/OGC/1.3/CRS84> Point(-
            0.001475 51.477811)
      </GeometryExpression>
      <GeometryDataType>
            ogc:WKTLiteral
      </GeometryDataType>
</Geometry>
```

GML

```
<Geometry>
	<GeometryExpression>
		<gml:Point
		srsName="http://www.opengis.net/def/crs/OGC/1.3/CRS84"><gml
		:coordinates>-0.001475, 51.477811</gml:coordinates>
		</gml:Point>
	</GeometryExpression>
	<GeometryDataType>
		ogc:GMLLiteral
	</GeometryDataType>
</Geometry>
```

URI reference (GeoHash)

```
<Geometry>
	<GeometryExpression>
		http://geohash.org/gcpuzgnzvxkp
	</GeometryExpression>
	<GeometryDataType>
		GeoHash
	</GeometryDataType>
</Geometry>
```

4. UNDEFINED CLASSES

The model shows two classes that are not yet defined. The working group strongly recommends that a new working group is formed to define them

4.1 CHANGE

Any record can only be seen as a snapshot of the data available to the publisher at a particular point in time. This is true for data related to all areas, not just people, places and businesses. The Change class is envisaged as one that captures any change in a data point. This might be a change in address, a change of name or any other update. It is likely that such a record would make it clear what data element was changed as well as when, why and by whom. This area - versioning and provenance - is a complex one that the working group hopes will be tackled by a subsequent activity as a matter of high priority.

The envisioned Change class entails a relationship:

Relationship	Abstract Data Type
change	Change

The change relationship links a Class to a record of change in a data field associated with the Class. For example, a change of name or a person's their death. It is assumed that an inverse relationship will also be defined.

4.2 LICENCE

In the context of a business register, the term licence refers to a permit to carry out a particular activity such as banking or mining. The Licence class is therefore envisaged as being similar to the Identifier Class in which a specific body is registered with a recognised authority. As with the Change class, the working group hopes that near future work can be done to ensure that such a class is well defined.

Two inverse relationships are entailed by the Licence Class:

Relationship	Abstract Data Type
holdsLicence	Licence

Relationship	Abstract Data Type
heldBy	Legal Entity

5. DATA TYPES

5.1 DATETIME

Dates recorded within public sector data sets exist in many different formats. In order to make those dates interoperable, it is important that they should be represented in a regular manner wherever possible. The most widely used date formats are those that conform to ISO 8601:2004 [ISO 8601] of which the relevant XML Datatypes [XSD] are a subset. Dates should be recorded as formatted strings that are conformant with that standard with the relevant data type declaration.

Known dates are recorded as literals in the form `yyyy-mm-dd` and then typed as `xsd:date`.

Where the full date and time are both known, and this can be important in records of death for example, use the `xsd:dateTime` format of `yyyy-mm-ddThh:mm:ss` with or without timezone information.

Where just a year is known use `xsd:gYear` (yyyy) and where just the month is known use xsd:`gYearMonth` (`yyyy-mm`). The 'g' is for Gregorian. Non-Gregorian dates are not covered by XSD.

For emphasis, the datatype should be explicitly stated so that software can correctly interpret the data. If a date is provided in a string that is conformant with a standard other than XSD, and that is not covered by XSD, then it should be typed accordingly. It should also be noted that the UK Government provides a URI scheme for time intervals that may be useful [DGUT] and that 8601:2004 itself allows intervals to be defined.

In some public sector data sets, full dates are not always known. For example, dates such as "some time in the 1920s" or "between 1925 and 1932" are not uncommon. There is no agreed formal way to record such uncertainty dates and therefore an un-typed plain string should be used. In other words, a string like "between 1925 and 1932" cannot be improved upon. To use an interval as the value for, say, a date of birth, is incorrect, The date of birth is not the time interval that began in 1925 and ended in 1932, rather, it was an instant some time during that period. There is an important difference between declaring that an event occurred throughout a time interval and that it occurred at an unknown instant within a time interval.

Some data sets use other methods to express uncertainty such as using false dates. Sometimes '00' or '??' is used to indicate uncertainty so that "August 1986" might be recorded as either 1986-00-00 or 1986-00-??. This practice is unnecessary and **strongly discouraged**, particularly if such dates are typed as XSD dates. If a string such as 1986-08-00 is fed to software expecting an xsd:date then it will either reject the data, in which case data is lost, or it might convert it to the nearest actual date 1986-08-01. In the latter case, the accurate yet imprecise date of "August 1986" has now been converted to a more precise date of 1st August 1986 that may be anything up to 30 days wide of the truth.

In summary:

- use the appropriate XSD data type wherever possible (dateTime, date, gYearMonth, gYear);
- where data is missing, don't invent it or try to fool the system - just give the date as an un-typed string.

5.2 TEXT

The text data type is a combination of a string and a language identifier. It is useful for names and descriptions that are available in multiple languages. Where this is so, each version of the data should be included and each one associated with the relevant language identifier. RFC 3066 [RFC 3066] provides a commonly used set of identifiers for natural languages. This is the set recognised by UN/CEFACT and XML Schema.

Languages are represented by two character codes, optionally followed by a locale definition such as "de" meaning German and "de-at" meaning "German as spoken in Austria."

XML Example:

```
<Location>
  <geographicName xml:lang="en">London</geoGraphicName>
  <geographicName xml:lang="fr">Londres</geoGraphicName>
</Location>
```

RDF Example:

```
[] a locn:Location ;
  locn:geographicName "London"@en ;
  locn:geographicName "Londres"@fr .
```

5.3 CODE

Interoperability between data sets is increased dramatically when terms from controlled vocabularies are used in favour of free text. The conceptual Code data type is used wherever one or more controlled vocabularies are known to exist for a particular domain of interest. It is not the job of Core Vocabularies to mandate which controlled vocabularies are used but we offer some guidance on how to use them.

It is important not only to use a term from a controlled vocabulary correctly (ideally a user interface would allow a term to be selected from a list rather than entered manually) but also to identify the source of the term, i.e. to use or, where necessary create, a datatype that is unambiguously associated with the term.

How this is done depends on the controlled vocabulary in question and how it is published.

5.3.1 XML Representation

To increase interoperability, the Core Data Type ‘Code’ of the UN/CEFACT Core Component Library [UN/CEFACT] can be used to provide metadata to the origin of the code. Where the publisher of the controlled vocabulary does not provide an XML schema it is necessary to use the Code datatype. For example, the following encodes the company type of " C18.01.02 - Other printing" according to Rev 2 of the European Commission's NACE codes.

```
<CompanyTypeCvbusinessCode>
            <Content>C18.01.02 - Other printing</Content>
            <List>NACE</List>
            <ListAgency>European Commission</ListAgency>
            <ListVersion> Rev 2</ListVersion>
<CompanyTypeCvbusinessCode>
```

5.3.2 RDF Representation

The Code datatype is best encoded in RDF as a SKOS Concept [SKOS]. Some controlled vocabularies are already published as such concept schemes. For example, to encode the female gender (section 3.1.6) using the SDMX code list, it is possible simply to provide http://purl.org/linked-data/sdmx/2009/code#sex-F as the venue of the gender property.

Where a vocabulary is not already available as a SKOS concept scheme, best practice is to create one as part of the data set (or better still, use someone else's encoding of it). For example, Jeni Tennison encoded the Observation Status codes from the SDMX Code lists [SDMX A2] for use by the UK Government following this approach [UKCODE].

Using this approach we might create a SKOS Concept scheme based on the ISO 5218 code as shown in the following Turtle:

```
@prefix rdf:     <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs:    <http://www.w3.org/2000/01/rdf-schema#> .
@prefix skos:    <http://www.w3.org/2004/02/skos/core#> .
@prefix schema:  <http://schema.org/> .
@prefix dcterms: <http://purl.org/dc/terms/> .

# First define a class that represents any value within the code list

<http://example.com/def/iso5218/sex>
  rdf:type <http://www.w3.org/2000/01/rdf-schema#Class> ;
  rdfs:label "Sex" ;
  rdfs:comment "This is the class of human sexes as defined by
  ISO 5218:2004" ;
```

```
10   rdfs:subClassOf <http://www.w3.org/2004/02/skos/core#Concept> .

# Now define the concept scheme itself.
# Notice that the URI for each term in the vocabulary is listed
# as a top concept. This is a 'flat' scheme. Use the full power
# of SKOS to describe hierarchical schemes. Other appropriate
# terms can be used to add further annotation about the concept scheme
# if necessary (see lines 24 and 25).

11 <http://example.com/id/iso5218/>
12   rdf:type <http://www.w3.org/2004/02/skos/core#ConceptScheme> ;
13   skos:prefLabel "Codes for the representation of human sexes"@en ;
14   skos:prefLabel "Codes de représentation des sexes humains"@fr ;
15   skos:notation "ISO/IEC 5218:2004" ;
16   skos:note "ISO 5218 specifies a uniform representation of human
17   sexes for the interchange of information" ;
18   skos:definition
19
<http://standards.iso.org/ittf/PubliclyAvailableStandards/c036266_ISO_
IEC_5218_2004(E_F).zip> ;
20   skos:hasTopConcept <http://example.com/id/iso5218/0> ;
21   skos:hasTopConcept <http://example.com/id/iso5218/1> ;
22   skos:hasTopConcept <http://example.com/id/iso5218/2> ;
23   skos:hasTopConcept <http://example.com/id/iso5218/9> ;
24   schema:version "2004" ;
25   dcterms:creator [dcterms:title "ISO"] .

# Now provide more detail about erach of those concepts

26   <http://example.com/iso5218/id/sex/0>
27   rdf:type <http://www.w3.org/2004/02/skos/core#Concept> ;
28   rdf:type <http://example.com/def/iso5218/sex> ;
29   skos:topConceptOf <http://example.com/id/iso5218/> ;
30   skos:prefLabel "Not known"@en ;
31   skos:prefLabel "Inconnu"@fr ;
# Note that labels can be provided in any number of languages
32   skos:notation "0" ; # The actual controlled vocabulary term is
                         # here, in this case, a single digit.
33   skos:definition "More detail of the vocabulary term if available
                   (it's not in the case of ISO 5218)" .

34   <http://example.com/iso5218/id/sex/1>
35   rdf:type <http://www.w3.org/2004/02/skos/core#Concept> ;
36   rdf:type <http://example.com/def/iso5218/sex> ;
37   skos:topConceptOf <http://example.com/id/iso5218/> ;
38   skos:prefLabel "Male"@en ;
```

```
39   skos:prefLabel "Masculin"@fr ;
40   skos:notation "1" .

41   <http://example.com/iso5218/id/sex/2>
42   rdf:type <http://www.w3.org/2004/02/skos/core#Concept> ;
43   rdf:type <http://example.com/def/iso5218/sex> ;
44   skos:topConceptOf <http://example.com/id/iso5218/> ;
45   skos:prefLabel "Feale"@en ;
46   skos:prefLabel "Féminin"@fr ;
47   skos:notation "2" .

48   <http://example.com/iso5218/id/sex/9>
49   rdf:type <http://www.w3.org/2004/02/skos/core#Concept> ;
50   rdf:type <http://example.com/def/iso5218/sex> ;
51   skos:topConceptOf <http://example.com/id/iso5218/> ;
52   skos:prefLabel "Not applicable"@en ;
53   skos:prefLabel "Sans objet"@fr ;
54   skos:notation "9" .
```

The following table shows how the terms used to create this concept scheme map to the Core Vocabularies Conceptual Model.

Conceptual Model	RDF term used
content	skos:prefLabel
list	skos:notation as a property of the concept scheme itself, in the example, this is ISA/IEC 5218:2004.
list agency	dcterms:creator (if the publisher has a URI, so much the better, if not, it may be necessary to create a blank node as shown in the example. The value of dcterms:creator should always be the URI of an Agent class, not a simple string or the URI of a Web site homepage etc, see line 25).
list version	schema:version (see line 24).

5.4 IDENTIFIER

The Identifier class is a critical aspect of the Business Core Vocabulary. As noted in section 3.1.12, it is also important in many data sets about individuals. In these cases and more, the identifier itself will be some sort of alpha-numeric string but that string only has meaning if it is contextualised. In the case of Legal Entities, the issuing of an identifier is a signal that legal entity status has been conferred on an organisation and it's important to know:

- the issuing agency;
- the date on which it was issued;
- the type of identifier issued (if the authority issues more than one type of identifier).

In general, identifiers are issued according to some sort of scheme and those schemes themselves may evolve over time and have version numbers. Section 5.8 of the UN/CEFACT Core Components Data Type Catalogue Version 3.1 defines a complex data type of Identifier with the following properties:

- content
- scheme identifier
- scheme identifier version
- scheme identifier agency

In terms of the conceptual model for the Core Vocabularies, we can interpret these as follows:

Conceptual Model	UN/CEFACT term
identifier	content
identifier type	scheme identifier
date of issue	N/A
issuing authority	scheme identifier agency
issuing authority URI	N/A
N/A	scheme identifier version

The Identifier data type is used in ADMS [ADMS] (which was developed in parallel to the Core Vocabularies) to record identifiers for semantic assets. This class is re-used for the Core Vocabularies as described below.

5.4.1 XML Representation

In an XML representation, the following snippet illustrates the use of the legal identifier. It is recommended to use URIs to encode the identifier type, and issuing authority ID.

```
<LegalIdentifierCvidentifier>
      <Identifier>String</Identifier>
      <IdentifierType>String</IdentifierType>
      <cbc:IssueDate>1967-08-13</cbc:IssueDate>
      <IssuingAuthority>String</IssuingAuthority>
      <IssuingAuthorityID>normalizedString</IssuingAuthorityID>
</LegalIdentifierCvidentifier>
```

5.4.2 RDF Representation

Taking the RDF Schema for ADMS [ADMSRDF] as a guide, the Identifier data type should be used as follows:

QName adms:Identifier

URI http://www.w3.org/ns/adms#Identifier

identifier	use skos:notation to provide the actual identifier (the Identifier class is effectively meaningless without this property and conceptually it is mandatory);
identifier type	use dcterms:type to provide an identifier for the type of identifier issued (the identifier scheme);
date of issue	use dcterms:created
issuing authority	use adms:schemeAgency
issuing authority	use dcterms:creator

If required, adms:schemeVersion can be used to identify the version of the identifier scheme.

6. PROPERTIES CONSIDERED AND EXCLUDED

6.1 RELATIONSHIPS

The Core Person conceptual model does not include the relationships that may exist between the described person and other entities such as family members, employers or property. Those relationships are clearly important in defining the role an individual plays in society and are often important in the definition of an individual's identity, however, they are not core to the identification of an individual in public sector data exchange.

6.2 PREFIXES, HONORARY TITLES AND SUFFIXES

Many family names include prefixes such as 'de' and 'von' and there is no simple way to decide whether or not such prefixes should be taken into account, for example, when sorting alphabetically. Herman Van Rompuy should be sorted under V whereas Ludwig van Beethoven should be sorted under B. Given the almost infinite variety of names and naming structures, there is no simple way to define the parts of names that will work across all possibilities. Such rules may be definable in a specific application or in a given country but that is out of scope for the Core Person vocabulary.

Therefore the working group resolved not to try and provide more granular properties for parts of names. It is perhaps noteworthy that the group very nearly resolved only to include the single property of Full Name with no further granularity. An excellent overview of different name conventions around the world is available from the W3C [PNAW].

6.3 ROLE

A specific relationship proposed as part of the Person Core Vocabulary was that of role. This could be used to link a record of any individual with the role they played, such as patient, company director or any other. It was narrowly decided not to include this relationship in this vocabulary. However, it is noted that in any given context, such a relationship can be expressed and the Person Core Vocabulary used. For example, a hospital patient record could point to a Person record in which case the 'patient' role is defined by the publisher of the patient record.

Feedback received on this point strongly suggests that role is important to the community, however, and so it is recommended that a mechanism for describing a person's role be included in future work. It is notable that both FOAF [FOAF] and the Organisation Ontology [ORG] include such a mechanism that also allows individuals to be grouped together and for that group to be assigned a role.

6.4 COUNTRY OF ORIGIN

The country or jurisdiction of origin is not included in the Business Core Vocabulary as this information will be obvious from the Identifier Class associated through the Legal Identifier.

6.5 HOLDS LICENCE & HELD BY

As noted in section 0, the business vocabulary sees a distinct need for a small vocabulary to describe licences issued to companies to permit them to carry out certain activities. The immediate use case is banking but there are many more highly regulated industries and the working group did not feel sufficiently knowledgeable to be able to capture the essential details. For example, the holder of a mining licence may be restricted to specified geographical areas, or times of day when they can operate; a waste disposal company may be restricted in the types of waste it can handle. A vocabulary is likely to need to support such details and this was deemed out of scope for the current work.

6.6 SPATIAL RELATIONSHIPS

The Location Working Group briefly considered including spatial relationships such as 'contains', 'borders' or 'overlaps' within this vocabulary. It was decided not to do so as where such properties are needed, publishers are likely to be familiar with more specialised vocabularies. However, it is worth noting that the Dublin Core term dcterms:relation or a sub property thereof may be appropriate.

7. USAGE GUIDELINES

The Core Vocabularies have been designed with maximum flexibility in mind with almost no restrictions on how any specific term can be used. In an RDF context, no new Domain and Range restrictions have been defined.

7.1 ALTERNATIVE ADDRESS REPRESENTATIONS (VCARD ETC.)

The INSPIRE method of representing addresses is very detailed, designed primarily for use in databases of addresses. Whilst data that is published in full conformance with the INSPIRE data structure can be made available using the Location Core Vocabulary the reverse is not true since the Core Vocabulary allows much greater flexibility. Many datasets that include address data as one piece of information about something else are likely to have that data in simpler formats. These might be tailored to the specific need of the dataset, follow a national norm, or make use of a standard like vCard [VC].

To provide maximum flexibility in the core vocabulary, whilst remaining interoperable with INSPIRE Address Guidelines (which EU Member States are obliged to use), the Location Core Vocabulary provides the extra property of full address (Section 3.5.2) and makes use of INSPIRE's addressID (section 3.5.12). As vCard is a massively used representation of addresses, we provide an example of how it can be fitted into the model but the same general method can be applied to any alternative address representation, such as OASIS xAL.

For our example we'll use the following (entirely fictitious) address:

15 Acacia Avenue
New Town
Edinburgh
UK
EH1 1AA

to which we'll assign an identifier of http://address.example/id/uk/eh11aa.

Using the Address Class defined in this specification the various properties and values would be as show in Table 1.

Property	Value
fullAddress	15 Acacia Avenue New Town Edinburgh UK EH1 1AA
locatorDesignator	15
thoroughfare	Acacia Avenue

addressArea	New Town
postName	Edinburgh
adminUnitL1	UK
postCode	EH1 1AA
addressID	http://address.example/id/uk/eh11aa

Table 1 An Edinburgh Address Represented using the Location Core Vocabulary

The mapping in Table 2 can be used to derive a vCard Address from this data. Note the inclusion of the addressID property which is not part of the vCard specification but is necessary so that the two representations can be associated.

As is clear, in this mapping it is really only the separation of the thoroughfare name from the locator properties that marks a difference beyond the particular terms used. It should be noted, however, that other mappings are possible and may be more appropriate in different contexts. The INSPIRE and vCard models are different, were designed for different purposes, and there are likely to be edge cases where the mapping doesn't work effectively. It is also notable that INSPIRE uses UUIDs for its address identifiers, not URIs, and so for full INSPIRE conformance, the UUID for a specific address should be used.

Mapping from granular data to the more coarse representation is relatively easy. However, the reverse is not possible in an automated way *unless* there are clear parsing rules which is unlikely. Therefore in situations where the available address data does not include, for example, separate locator designator and thoroughfare information, publishers should simply provide the fullAddress and addressID properties and then use the ID to associate the address with the vCard representation or whatever other representation is appropriate.

vCard	Location Core Vocabulary
street-address	concatenated string comprising: locatorDesignator locatorName thoroughfare
extended-address	addressArea
locality	postName
region	adminUnitL2
country-name	adminUnitL1
postal-code	postcode
label	fullAddress
addressed	http://address.example/id/uk/eh11aa

Table 2 An Edinburgh Address represented in vCard

7.2 BUSINESS VOCABULARY EXAMPLE

The RDF Schema package includes an example of a company described using the Business and Location vocabularies (in Turtle). This includes:

- a description of the legal entity;
- a legal identifier (i.e. details of the registration with the authority that conferred legal status);
- two further identifiers (Open Corporates and the European Business Register) ;
- the company's registered address (INSPIRE-conformant);
- a extract from a possible encoding of the relevant NACE codes using SKOS.

The example is reproduced as Appendix A.

7.3 PERSON VOCABULARY EXAMPLE

Appendix B provides an example of a description of a (real) person in RDF (Turtle). It re-uses the ISO 5218 example used in the Code data type example (5.3).

8. RDF AND XML SCHEMAS FOR THE CORE VOCABULARIES

Serialisations of each of the Core Vocabularies in both XML and RDF are available on the Joinup Platform.

8.1 RDF SCHEMA

The RDF schemas for the Core Vocabularies are provided separately and qre serialised in RDF/XML with an XSLT that renders it into an HTML page for human readers. The majority of terms used are taken from existing vocabularies, notably Dublin Core and FOAF. Newly minted terms are specific to the concepts of the Core Vocabularies.

Domain and range restrictions have not been defined for terms borrowed from existing vocabularies and have only been defined sparingly for the Core Vocabularies.

8.2 XML SCHEMA

The backbone of the Core Vocabularies in XML is a Common Library of information elements provided by the library of the Universal Business Language (UBL)[2]. The philosophy behind this design is to achieve reusability of information elements defined by the Core Component Technical Specification (CCTS)[3] of UN/CEFACT[4] (the basis of UBL).
(See diagram on next page)

[2] OASIS. Universal Business Language v2.1. http://docs.oasis-open.org/ubl/UBL-2.1.pdf

[3] United Nations Centre for Trade Facilitation and Electronic Business. Core Components Technical Specification Version 3.0 http://www.unece.org/fileadmin/DAM/cefact/codesfortrade/CCTS/CCTS-Version3.pdf

[4] UNECE United Nations Economic Commission for Europe. UN Centre for Trade Facilitation and E-business (UN/CEFACT). http://www.unece.org/cefact/index.html

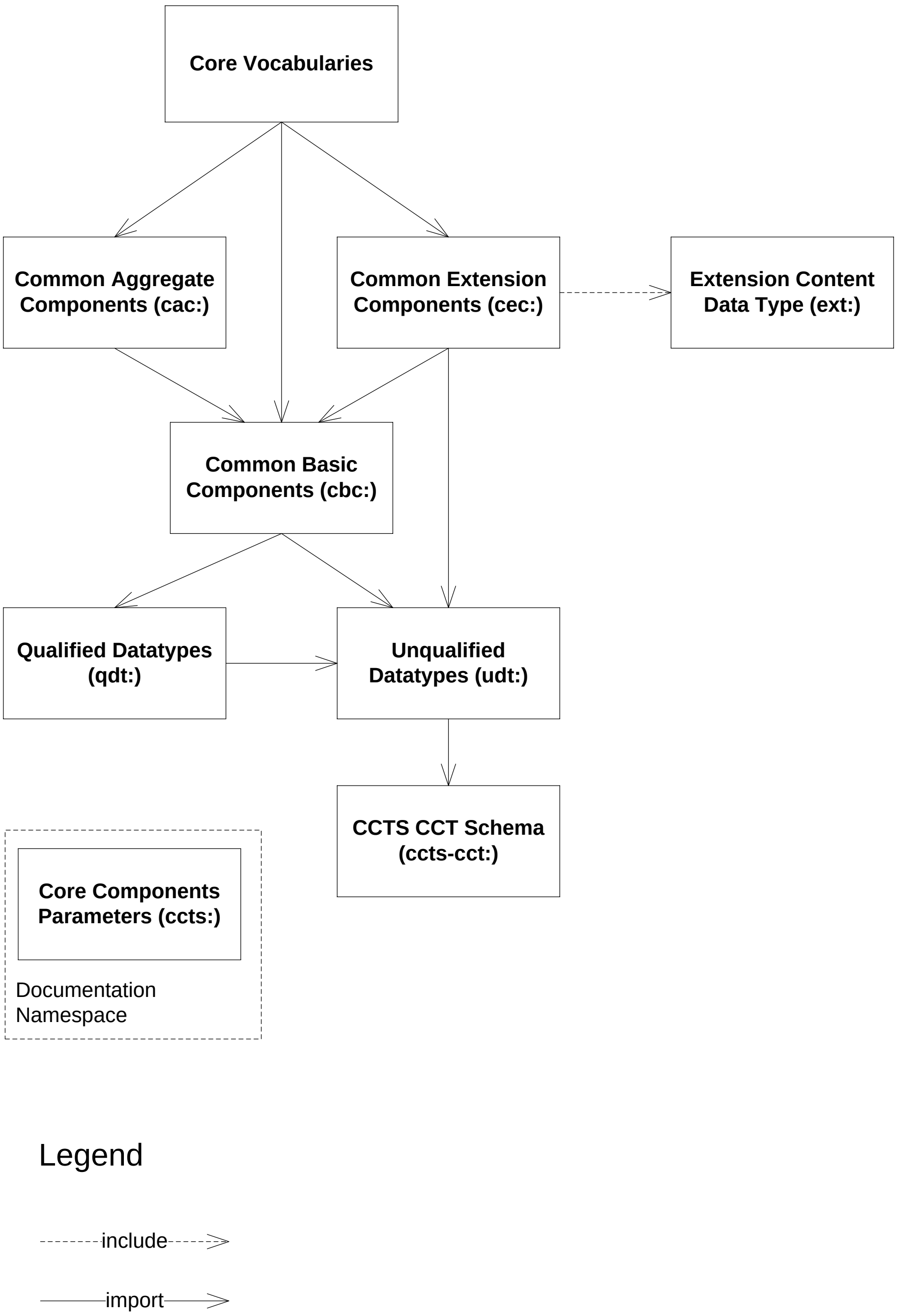

Figure 2 Core Vocabularies XML Schema and reuse of UBL and CCTS

Reused schemas:

CommonBasicComponents[5]

The CommonBasicComponents schema defines the global Basic Business Information Entities (BBIEs) that are used throughout UBL, serving, in effect, as a “global BBIE type

[5] http://docs.oasis-open.org/ubl/prd1-UBL-2.1/xsd/common/UBL-CommonBasicComponents-2.1.xsd

database" for constructing documents. BBIEs are the "leaf nodes" of UBL documents, corresponding to individual data fields in traditional printed business forms.

CommonAggregateComponents[6]

The CommonAggregateComponents schema defines the Aggregate Business Information Entities (ABIEs) that are used throughout UBL, serving, in effect, as an "ABIE type database" for constructing the main documents.

UnqualifiedDataTypes[7]

This schema defines Unqualified Data Types for BBIE definition. These types are derived from the Core Component Types in CCTS_CCT_SchemaModule[8]. Where an unqualified type is not based solely on an XSD data type, all CCTS supplementary components are made available in the UBL Unqualified Data Type from the CCTS_CCT_SchemaModule.

QualifiedDataTypes[9]

CCTS permits the definition of Qualified Datatypes as derivations from CCTS specified Unqualified Datatypes. In UBL 2.1, all data type qualifications are expressed in the file cva/UBL-DefaultDTQ-2.1.cva[10].The UBL-QualifiedDataTypes-2.1.xsd file in the UBL 2.1 release is included among the schema modules imported by the Common Library and all document-level schema fragments in order to be consistent with the relationship of types

---

[6] http://docs.oasis-open.org/ubl/prd1-UBL-2.1/xsd/common/UBL-CommonAggregateComponents-2.1.xsd
[7] http://docs.oasis-open.org/ubl/prd1-UBL-2.1/xsd/common/UBL-UnqualifiedDataTypes-2.1.xsd
[8] http://docs.oasis-open.org/ubl/prd1-UBL-2.1/xsd/common/CCTS_CCT_SchemaModule-2.1.xsd
[9] http://docs.oasis-open.org/ubl/prd1-UBL-2.1/xsd/common/UBL-QualifiedDataTypes-2.1.xsd
[10] http://docs.oasis-open.org/ubl/prd2-UBL-2.1/cva/UBL-DefaultDTQ-2.1.cva

9. BACKGROUND AND OBJECTIVES

As noted in the introduction, these Core Vocabularies are one of a set produced under Action 1.1 of the ISA Programme [A1.1]. Further background is available in "e-Government Core Vocabularies [EGOV-CV] which offers an overview and context for the work.

The natural course of action for any practitioner or team given the task of recording information about a natural person is to write a list of the data elements they need (or already have) and to work within the specific context of their project. This often works in that it demonstrably meets the project's needs. The problems only arise when one team wants to exchange data with another. It's at that point that the choice of, say, 'first name' over 'given name' and 'surname' over 'family name' becomes an obstacle. Such terms are well defined in a variety of vocabularies and their use cannot be regarded as 'wrong', however, it's easy for simple choices to lead to unintended difficulties further down the road.

The aim of providing Core Vocabularies via the Joinup Platform is not to force teams to use a particular set of terms, or to require the re-engineering of data sets to use them (which can be prohibitively expensive). Rather the aim is to make it easy to see and use the terms that crop up across multiple domains; terms that, when used by public sector agencies, will make data more interoperable.

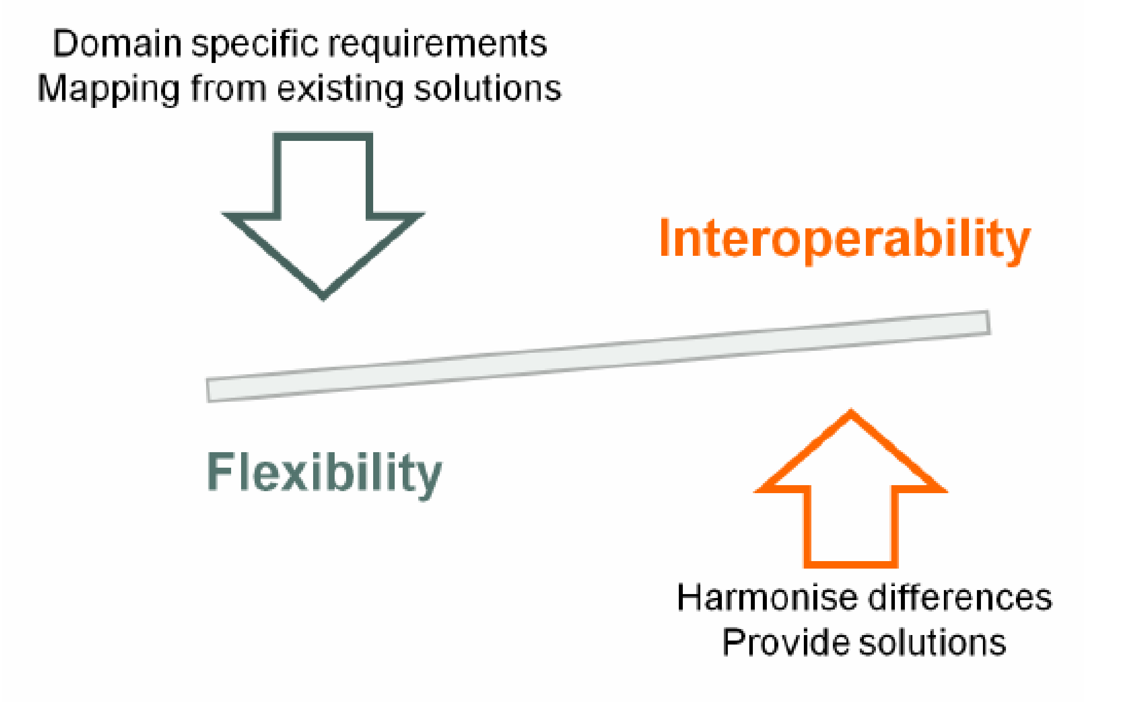

Figure 3 The struggle between enabling interoperability and giving flexibility

Identifiers are a case in point. In an international context, a person's passport number is likely to be critically important. This is not so within a university where the likelihood is that an in-house identifier will be assigned. Rather than one agency defining a term for 'passport number' and another for 'student number', both can use the core vocabulary term 'identifier', preferably with some additional contextual information. As Figure 3 illustrates, there is a balance to be struck between flexibility and interoperability.

10. APPROACH & COMMUNITY

Early work under the ISA Programme, known as SEMIC, produced an initial Core Person vocabulary specification [SCP]. This specification draws directly on that work and extends it slightly for the Person Class (all the terms in the SEMIC Core Person Vocabulary exist within this update). The Business and Location vocabularies are 'new' however, as is clear, they both draw heavily on existing work, particularly the Location vocabulary which is driven by work on the INSPIRE directive.

The process and methodology followed in the development is set out in detail in the Process and Methodology for Developing Core Vocabularies [PMDCV].

Specific acknowledgement is due to:

- Joseph Azzopardi, Malta Information Technology Agency
- Morten Borrebæk, Norwegian Mapping Authority
- Ignacio Boixo, Banco de España
- Serena Coetzee, ISO/TC 211
- Stijn Goedertier, PwC EU Services
- Christophe Guéret, Vrije Univerisiteit Amsterdam
- Martin Herzog, Init
- Anja Hopfstock, Federal Agency for Cartography and Geodesy (BKG), Germany
- Anne Gro Hustoft, Agency for Public Management and eGovernment, Norway
- Michaela Elisa Jackson, CSI-Piemonte
- Morten Lind, Danish Enterprise and Construction Authority
- Michael Lutz, Joint Research Centre
- Piotr Madziar, DG MARKT
- Andras Micsik, MTA SZTAKI
- Aine Mitchell, EBR
- Andrea Perego, Joint Research Centre
- Greg Potterton, Eurojust
- Raj Singh, Open Geospatial
- Sebastian Sklarß, Init.
- Paul Smits, Joint Research Centre
- Chris Taggart, Open Corporates
- Thomas Verdin, XBRL Europe
- Luk Vervenne, Synergetics, Belgium
- Philippe Vlerick, EC, DG Justice
- Neven Vrček, University of Zagreb
- Rob Walker, ISO/TC 211
- Stuart Williams, Epimorphics
- Peter Winstanley, Scottish Government

11. CHANGE CONTROL

The Core Vocabularies are published by the ISA Programme. Review comments and requests for changes can be made via the mailing list which is archived at http://joinup.ec.europa.eu/mailman/private/isa-cv-person-wg/.

12. REFERENCES

[A1.1] Action 1.1 Improving semantic interoperability in European eGovernment systems http://ec.europa.eu/isa/actions/01-trusted-information-exchange/1-1action_en.htm

[ADMS] The Asset Description Metadata Standard, http://joinup.ec.europa.eu/asset/adms/description

[ADMSRDF] See https://joinup.ec.europa.eu/system/files/project/ADMS_RDF_Schema_v1.00.zip

[RADion] RADion - a Repository, Asset, Distribution framework, see http://joinup.ec.europa.eu/asset/radion/home

[DBpedia] DBpedia is a community effort to extract structured information from Wikipedia and to make this information available on the Web. http://dbpedia.org/

[DC] DCMI Metadata Terms, Dublin Core Metadata Initiative. http://dublincore.org/documents/dcmi-terms/

[DGUT] data.gov.uk Time Intervals. Linked data for every time interval and instant into the past and future, from years down to seconds. This is an infinite set of linked data. It includes government years and properly handles the transition to the Gregorian calendar within the UK. http://thedatahub.org/dataset/data-gov-uk-time-intervals

[EGOV-CV] e-Government Core Vocabularies: The SEMIC.EU approach. Retrieved from European Commission - Directorate-General Informatics: http://www.semic.eu/semic/view/documents/egov-core-vocabularies.pdf

[FOAF] Friend of a Friend http://xmlns.com/foaf/spec/

[GML] Geography Markup Language, Open Geospatial Consortium Standard. http://www.opengeospatial.org/standards/gml

[INADR] D2.8.I.5 INSPIRE Data Specification on Addresses – Guidelines. Infrastructure for Spatial Information in Europe (INSPIRE). http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_AD_v3.0.1.pdf

[INGN] D2.8.I.3 INSPIRE Data Specification on Geographical Names – Guidelines. http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_GN_v3.0.1.pdf

[ISA] Interoperability Solutions for European Public Administrations, http://ec.europa.eu/isa/

[ISIC4] International Standard Industrial Classification of All Economic Activities, Rev.4, United Nations Statistics Division. See http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=27

[ISO 3166-1] Codes for the representation of names of countries and their subdivisions – Part 1: Country codes. http://www.iso.org/iso/iso_3166_code_lists

[ISO 3166-3] Codes for the representation of names of countries and their subdivisions – Part 3: Code for formerly used names of countries (most easily accessed via Wikipedia at http://en.wikipedia.org/wiki/ISO_3166-3)

[ISO 5218] Codes for the representation of human sexes http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=36266

[ISO 8601] Data elements and interchange formats -- Information interchange -- Representation of dates and times, ISO 8601:2004. http://www.iso.org/iso/catalogue_detail?csnumber=40874

[ISO 19125] ISO 19125-1:2004 Geographic information -- Simple feature access -- Part 1: Common architecture See http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=40114)

[ISO 19112] Geographic information -- Spatial referencing by geographic identifiers (ISO 19112:2003) http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=26017

[JOINUP] The Joinup Platform is operated by the European Commission designed to enable the sharing and reuse open-source software, semantic assets and other interoperability solutions for public administrations. See http://joinup.ec.europa.eu/

[NACE] Statistical Classification of Economic Activities in the European Community. Eurostat. http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL&StrNom=NACE_REV2&StrLanguageCode=EN

[ORG] An organization ontology, Dave Reynolds/W3C http://www.w3.org/TR/vocab-org/

[PMDCV] Process and Methodology for Developing Core Vocabularies, 22 November 2011. https://joinup.ec.europa.eu/elibrary/document/isa-deliverable-process-and-methodology-developing-core-vocabularies

[PNAW] Personal names around the world, Richard Ishida, W3C. http://www.w3.org/International/questions/qa-personal-names

[RFC 3066] H. Alvestrand, ed. RFC 3066: Tags for the Identification of Languages 1995. Available at: http://www.ietf.org/rfc/rfc3066.txt

[SCLS] Eurostat Standard Code Lists are available through the RAMON portal (http://ec.europa.eu/eurostat/ramon/).

[SDMX A2] SDMX Content-Oriented Guidelines, Annex 2, Corss Domain Code Lists. http://sdmx.org/wp-content/uploads/2009/01/02_sdmx_cog_annex_2_cl_2009.pdf

[SCP] SEMIC Core Person. Available via http://joinup.ec.europa.eu/asset/semic-eu_core_person_17/home

[SDMX CDL] Content-Oriented Guidelines, Annex 2: Cross-Domain Code Lists http://sdmx.org/wp-content/uploads/2009/01/02_sdmx_cog_annex_2_cl_2009.pdf.

[SKOS] SKOS Simple Knowledge Organization System, Reference. Miles, A, Bechhofer, S, W3C Recommendation 18 August 2009. http://www.w3.org/TR/skos-reference/

[TOGM] Towards Open Government Metadata, Vassilios Peristeras, DG DIGIT, ISA Unit, September 2011 https://joinup.ec.europa.eu/sites/default/files/towards_open_government_metadata_0.pdf

[TURTLE] Terse RDF Triple Language, W3C http://www.w3.org/TR/turtle/

[UKCODE] Observation Status concept scheme. http://www.jenitennison.com/blog/files/codelists.ttl

[UN/CEFACT] UN Centre for Trade Facilitation and e-Business http://www.unece.org/cefact/

The UN/CEFACT Core Components Library (CCL) can be found at http://www.unece.org/cefact/codesfortrade/unccl/CCL_index.html

[UPU] Universal Postal Union (UPU), "International Postal Address Components and Templates", UPS SB42-4, July 2004

[VC] vCard Specifications, Internet Mail Consortium. See http://www.imc.org/pdi/

[W3CDTF] Date and Time Formats. W3C Note 15 September 1997. http://www.w3.org/TR/NOTE-datetime

[W3CWGS] WGS84 Geo Positioning: an RDF vocabulary, W3C Note 2003 http://www.w3.org/2003/01/geo/wgs84_pos

[WKT] OGC GeoSPARQL - A Geographic Query Language for RDF Data Under development at the Open Geospatial Consortium (http://www.opengeospatial.org)

[XAL] OASIS xAL Standard v2.0. See http://www.oasis-open.org/committees/ciq/download.html

[xEBR] XBRL Europe http://www.xbrl.org/

[XSD] XML Schema Part 2: Datatypes Second Edition. W3C Recommendation 28 October 2004. http://www.w3.org/TR/xmlschema-2/#date

13. THE CORE VOCABULARIES WORKING GROUP

Given name	Family name	Organisation	Country	Task Force
Segun	Alayande	SmartInformatics Limited - british airports	UK	Core Person, Core Location, Core Business
Miguel	Alvarez Rodríguez	Ministerio de Política Territorial y Administración Pública	ES	Core Person
Adam	Arndt	The Danish Agency for Digitization	DK	Core Person, Core Location, Core Business
Joseph	Azzopardi	Malta Information Technology Agency	MT	Core Person, Core Location, Core Business
Christina	Bapst	NIEM - National Information Exchange Models	US	Core Person, Core Location, Core Business
Peter	Bjerghave	Danish National Police	DK	Core Person
Lyubomir	Blagoev	USW Ltd	BG	Core Person, Core Location, Core Business
Ignacio	Boixo	Eurofilling.info	ES	Core Business
Morten	Borrebæk	Norwegian Mapping Authority	NO	Core Location
Dan	Brickley	Vrije Universiteit	NL	Core Person, Core Location, Core Business
Heath	Bunting	Irational	UK	Core Person, Core Location, Core Business
Peter	Burian	DIGIT	EU	Core Business, Core Location, Core Person

Giorgio	Cangioli	Dedalus S.p.A	IT	Core Person
Serena	Coetzee	University of Pretoria	ZA	Core Location
Uwe	Ewald	International Justice Analysis Forum	DE	Core Person
Paul	Farrell	EBR		Core Business
Corine	Frappart	Publications Office of the European Union	EU	Core Person, Core Location, Core Business
Muriel	Foulonneau	Tudor Public Research Centre	LU	Core Person
Marc	Fresko	Inforesight	UK	Core Person, Core Business
Andreas	Gehlert	Federal Ministry of Interior - Germany	DE	Core Person, Core Location, Core Business
Giorgios	Georgiannakis	DG SANCO	BE	Core Person, Core Location, Core Business
Stijn	Goedertier	PwC Belgium	BE	Core Person, Core Location, Core Business
Christophe	Guéret	Vrije Universiteit	NL	Core Person, Core Location
Anja	Hopfstock	Federal Agency for Cartography and Geodesy (BKG), Germany	DE	Core Location
Anne Gro	Hustoft	Agency for Public Management and eGovernment	NO	Core Person
Andreas	Illert	BKG Bund.de	DE	Core Location

Michaela Elisa	Jackson	CSI-Piemonte - Governo e Gestione	IT	Core Person, Core Location
Brigitte	Joerg	DFKI	DE	Core Person, Core Business
Anne	Kauhanen-Simanainen	Ministry of Finance	FI	Core Person, Core Location, Core Business
Thomas	Knape	Data Fusion/IEEE	IE	Core Person, Core Location, Core Business
Pim	Keizer	Dutch Ministry of Justice	NL	Core Business
Madeleine	Kiss	Publications Office of the European Union	EU	Core Person, Core Location, Core Business
Morten	Lind	Danish Enterprise and Construction Authority	DK	Core Location
Nikos	Loutas	DERI	IE	Core Person, Core Location, Core Business
Michael	Lutz	JRC	IT	Core Location
Matthias	Lüttgert	ENDA	DE	Core Business, Core Location
Martin	Määrits	Centre of Registers and Information Systems	EE	Core Business
Piotr	Madziar	DG MARKT	BE	Core Business
Charles-Henri	Menseau	Département Titres Identités	FR	Core Person
Andras	Micsik	SZTAKI	HU	Core Person, Core Location

Aine	Mitchell	Enterprise Registry Solutions / EBR	IE	Core Business
Natalie	Nickel	e-Codex	DE	Core Person
Priit	Parmakson	Estonian Information Systems Authority	EE	Core Person
Andrea	Perego	JRC	IT	Core Location
Vassilios	Peristeras	DG DIGIT	GR	Core Person, Core Location, Core Business
Piotr	Piotrowski	Universal Postal Union; Address Data Management		Core Location
Greg	Potterton	Eurojust	BE	Core Person, Core Location, Core Business
Peter	Reichstädter	Bundeskanzleramt / Repository owner	AT	
Børge Krogh	Samuelsen	KL - Denmark	DK	Core Person
Peter	Schmitz	Publications Office of the European Union	EU	Core Person, Core Business, Core Location
Raj	Singh	Open Geospatial Consortium	US	Core Location
Sebastian	Sklarß	]init[	DE	Core Person, Core Business, Core Location
Ronald	Smit	e-Codez	NL	Core Person
Dan	Smith	LinkedGov	UK	Core Person, Core Business, Core Location

Paul	Smits	JRC	IT	Core Location
Ernst	Steigenga	e-Codex	NL	Core Person
Chris	Taggart	Open Corporates	UK	Core Business
Athina	Trakas	OGC - Open Geospacial Consortium	GR	Core Location
Maris	Tudre	Centre of Registers and Information Systems	EE	Core Business
Thomas	Verdin	XBRL	BE	Core Person, Core Business
Luk	Vervenne	Synergetics	BE	Core Person
Philippe	Vlerick	DG JUST	BE	Core Person
Neven	Vrček	University of Zagreb	HR	Core Person, Core Business
Rob	Walker	Independent Consultant	UK	Core Location
Andreas	Weller	European Banking Authority EBA	LU	Core Business
Stuart	Williams	Epimorphics	UK	Core Location
Peter	Winstanley	ISIS:KIRM - The Scottish Government	UK	Core Person, Core Location, Core Business
Glyn	Wintle	LinkedGov	UK	Core Person, Core Location, Core Business

The SEMIC Team:

- Phil Archer, W3C (Core Vocabularies Editor)
- Makx Dekkers, AMI Consult
- Michiel De Keyzer, PwC
- Débora Di Giacomo, PwC
- João Rodrigues Frade, PwC
- Stijn Goedertier, PwC
- Thomas Rössler, W3C
- Niels Van Hee, PwC
- Rigo Wenning, W3C

14. APPENDIX A – NAMESPACES

On publication of this document by the ISA Programme, the Core Vocabularies will become an input to the W3C Government Linked Data Working Group [GLD]. Through that process the vocabulary will be hosted on the w3.org domain. The namespaces and suggested prefixes for the three core vocabularies are thus:

Prefix	Namespace
`person`	`http://w3.org/ns/person#`
`legal`	`http://w3.org/ns/legal#`
`locn`	`http://w3.org/ns/locn#`

15. APPENDIX B – CORE BUSINESS EXAMPLE

This is a listing of the example of a description of a business created using the relevant Core Vocabularies. It uses a similar construct for the encoding of the NACE codes as described in section 5.3.

```
@prefix rdf:     <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs:    <http://www.w3.org/2000/01/rdf-schema#> .
@prefix skos:    <http://www.w3.org/2004/02/skos/core#> .
@prefix dcterms: <http://purl.org/dc/terms/> .
@prefix legal:   <http://www.example.org/ns/legal#> .
@prefix locn:    <http://www.example.org/ns/locn#> .
@prefix adms:    <http://www.w3.org/ns/adms#> .
@prefix foaf:    <http://xmlns.com/foaf/0.1/> .
@prefix xsd:     <http://www.w3.org/2001/XMLSchema> .
@prefix schema:  <http://schema.org/> .

# This is an example of a (real) legal entity and how it is
# described using the Business Core Vocabulary. We use example.com
# to represent the body that publishes the data.

<http://data.companieshouse.gov.uk/id/company/04285910>
  a legal:LegalEntity ;
  legal:legalName "Apple Binding Ltd" ;
  legal:companyStatus <http://example.com/id/status/NormalActivity> ;
  legal:companyType <http://example.com/id/type/Plc> ;
  legal:companyActivity <http://example.com/id/NACE/2/C/18/01/02> ;
  legal:companyActivity <http://example.com/id/NACE/2/C/18/01/04> ;
  # Note the use of legalIdentifier on the next line
  legal:legalIdentifier <http://example.com/id/li04285910> ;
  # Other identifiers are often very useful but they do
  # not confer legal entity status.
  legal:identifier <http://example.com/id/oc04285910> ;
  legal:identifier <http://example.com/id/ebr04285910> ;
  legal:registeredAddress <http://example.com/id/ra04285910> .

# The actual registration
<http://example.com/id/li04285910> a adms:Identifier ;
  skos:notation "04285910" ;
  adms:schemaAgency "Companies House" ;
  dcterms:created "2001-09-12"^^xsd:date ;
  dcterms:creator <http://example.com/id/foaf/CompaniesHouse> .
```

```
# A supplementary identifier (Open Corporates)
<http://example.com/id/oc04285910> a adms:Identifier ;
  skos:notation "http://opencorporates.com/companies/gb/04285910" ;
  dcterms:created "2010-10-21T15:09:59Z"^^xsd:dateTime ;
  dcterms:modified "2012-04-26T15:16:44Z"^^xsd:dateTime ;
  dcterms:creator <http://example.com/id/foaf/OpenCorporates> .

# A further identifier (the REID number from the EBR)
<http://example.com/id/ebr04285910>a adms:Identifier ;
  skos:notation "GBCHC100.04285910-77" ;
  dcterms:created "2012-02-27"^^xsd:date ;
  dcterms:creator <http://example.com/id/foaf/EBR> .

# Some basic info about the ID-issuing authorities
<http://example.com/id/foaf/CompaniesHouse> a foaf:Organization ;
  foaf:name "Companies House" ;
  foaf:homepage <http://www.companieshouse.gov.uk/> .

<http://example.com/id/foaf/OpenCorporates> a foaf:Organization ;
  foaf:name "Open Corporates" ;
  foaf:homepage <http://opencorporates.com/> .

<http://example.com/id/foaf/EBR> a foaf:Organization ;
  foaf:name "European Business Register (EBR)" ;
  foaf:homepage <http://www.ebr.org//> .

# The registered address of the company (INSPIRE-conformant)
<http://example.com/id/ra04285910> a locn:Address ;
  locn:locatorDesignator "17" ;
  locn:thoroughfare "Riverside Ave West" ;
  locn:addressArea "Lawford" ;
  locn:postName "Manningtree" ;
  locn:adminUnitL2 "Essex" ;
  locn:adminUnitL1 "UK" ;
  locn:postCode "CO11 1UN" ;
  locn:fullAddress """17 Riverside Ave West
Lawford
Manningtree
Essex
UK
CO11 1UN""" .

# NACE Codes - here is a snapshop of a possible SKOS representation
# We'll just define the classes we need for this example here.
# Similar constructs can be used to represent controlled vocabularies
# for companyType and companyStatus but these are not included in this
```

```
# example. First define a class that represents any value within the
# code list

<http://example.com/def/NACE/code>
  rdf:type <http://www.w3.org/2000/01/rdf-schema#Class> ;
  skos:prefLabel "NACE Code" ;
  rdfs:comment "This is the class of NACE codes as defined by the
European Commission"@en ;
  rdfs:subClassOf <http://www.w3.org/2004/02/skos/core#Concept> .

# Now define the concept scheme itself.

<http://example.com/id/NACE/2/>
  rdf:type <http://www.w3.org/2004/02/skos/core#ConceptScheme> ;
  skos:prefLabel "General Name for Economic Activities in the European
Union"@en ;
  skos:prefLabel "Nomenclature Generale des Activites Economiques dans
I`Union Europeenne"@en ;
  skos:notation "NACE Revision 2" ;
  skos:note "The latest NACE codes (Revision 2) are based on the
Regulation (EC) No 1893/2006 of the European Parliament and of the
Council, establishing the statistical classification of economic
activities." ;
  schema:version "2008" ;
  dcterms:creator [dcterms:title "European Commission"] ;
  skos:definition
<http://ec.europa.eu/competition/mergers/cases/index/nace_all.html> ;
  skos:hasTopConcept <http://example.com/id/NACE/2/A> ;
  skos:hasTopConcept <http://example.com/id/NACE/2/B> ;
  skos:hasTopConcept <http://example.com/id/NACE/2/C> ;
  skos:hasTopConcept <http://example.com/id/NACE/2/D> ;
  skos:hasTopConcept <http://example.com/id/NACE/2/E> ;
  skos:hasTopConcept <http://example.com/id/NACE/2/F> ;
  skos:hasTopConcept <http://example.com/id/NACE/2/G> ;
  skos:hasTopConcept <http://example.com/id/NACE/2/H> ;
  skos:hasTopConcept <http://example.com/id/NACE/2/I> ;
  skos:hasTopConcept <http://example.com/id/NACE/2/J> ;
  skos:hasTopConcept <http://example.com/id/NACE/2/K> ;
  skos:hasTopConcept <http://example.com/id/NACE/2/L> ;
  skos:hasTopConcept <http://example.com/id/NACE/2/M> ;
  skos:hasTopConcept <http://example.com/id/NACE/2/N> ;
  skos:hasTopConcept <http://example.com/id/NACE/2/O> ;
  skos:hasTopConcept <http://example.com/id/NACE/2/P> ;
  skos:hasTopConcept <http://example.com/id/NACE/2/Q> ;
  skos:hasTopConcept <http://example.com/id/NACE/2/R> ;
  skos:hasTopConcept <http://example.com/id/NACE/2/S> ;
```

```
  skos:hasTopConcept <http://example.com/id/NACE/2/T> ;
  skos:hasTopConcept <http://example.com/id/NACE/2/U> .

# Now provide more detail about each of those concepts

  <http://example.com/id/NACE/2/C>
  rdf:type <http://www.w3.org/2004/02/skos/core#Concept> ;
  rdf:type <http://example.com/def/NACE/code> ;
  skos:topConceptOf <http://example.com/id/NACE/2/> ;
  skos:prefLabel "Section C : Manufacturing";
  skos:notation "C" .

  <http://example.com/id/NACE/2/C/18>
  rdf:type <http://www.w3.org/2004/02/skos/core#Concept> ;
  rdf:type <http://example.com/def/NACE/code> ;
  skos:broader  <http://example.com/id/NACE/2/C> ;
  skos:inScheme <http://example.com/id/NACE/2/> ;
  skos:prefLabel "C18 : Printing and reproduction of recorded media";
  skos:notation "C18" .

  <http://example.com/id/NACE/2/C/18/01>
  rdf:type <http://www.w3.org/2004/02/skos/core#Concept> ;
  rdf:type <http://example.com/def/NACE/code> ;
  skos:broader  <http://example.com/id/NACE/2/C/18> ;
  skos:inScheme <http://example.com/id/NACE/2/> ;
  skos:prefLabel "C18.01 : Printing and service activities related to
printing";
  skos:notation "C18.01" .

  <http://example.com/id/NACE/2/C/18/01/02>
  rdf:type <http://www.w3.org/2004/02/skos/core#Concept> ;
  rdf:type <http://example.com/def/NACE/code> ;
  skos:broader  <http://example.com/id/NACE/2/C/18/01> ;
  skos:inScheme <http://example.com/id/NACE/2/> ;
  skos:prefLabel "C18.01.02 : Other printing";
  skos:notation "C18.01.02" .

  <http://example.com/id/NACE/2/C/18/01/04>
  rdf:type <http://www.w3.org/2004/02/skos/core#Concept> ;
  rdf:type <http://example.com/def/NACE/code> ;
  skos:broader  <http://example.com/id/NACE/2/C/18/01> ;
  skos:inScheme <http://example.com/id/NACE/2/> ;
  skos:prefLabel "C18.01.04 : Binding and related services";
  skos:notation "C18.01.04" .
```

16. APPENDIX C – CORE PERSON EXAMPLE

This is a listing describing UK politician Paddy Ashdown using the Person Core Vocabulary. It is available as a separate file along with the RDF schema.

```
@prefix rdf:     <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs:    <http://www.w3.org/2000/01/rdf-schema#> .
@prefix skos:    <http://www.w3.org/2004/02/skos/core#> .
@prefix dcterms: <http://purl.org/dc/terms/> .
@prefix legal:   <http://www.example.org/ns/legal#> .
@prefix locn:    <http://www.example.org/ns/locn#> .
@prefix adms:    <http://www.w3.org/ns/adms#> .
@prefix person:  <http://www.w3.org/ns/person#> .
@prefix foaf:    <http://xmlns.com/foaf/0.1/> .
@prefix xsd:     <http://www.w3.org/2001/XMLSchema> .
@prefix schema:  <http://schema.org/> .

# This is an example of a (real) legal entity and how it is described
# using the Business Core Vocabulary. We use example.com to represent
# the body that publishes the data

<http://example.com/people/001> a person:Person ;
  foaf:name "Jeremy John Durham Ashdown, Baron Ashdown of Norton-sub-
Hamdon, GCMG, KBE, PC" ;
  foaf:givenName "Jeremy John Durham" ;
  foaf:familyName "Ashdown" ;
  person:alternativeName "Paddy Ashdown" ;
  person:alternativeName "Lord Ashdown" ;
  schema:gender <http://example.com/id/iso5218/1> ;
  person:birthName "Jeremy John Durham Ashdown" ;
  schema:birthDate "1941-02-27"^^xsd:date ;
  person:placeOfBirth [a locn:Location ; locn:geographicName "New
Delhi"@en ; locn:geographicIdentifier
<http://sws.geonames.org/1261481/>; ] ;
  person:countryOfBirth [a locn:Location ; locn:geographicName
"India"@en ; locn:geographicIdentifier
<http://dbpedia.org/resource/ISO_3166-1:IN>; ] ;
  person:citizenship <http://example.com/jurisdiction/id/UK> .

<http://example.com/jurisdiction/id/UK> a dcterms:Jurisdiction ;
  rdfs:label "UK" .
```