HORIZON 2020
ICT - INFORMATION AND COMMUNICATION TECHNOLOGIES

PROFIT
Promoting Financial Awareness and Stability

PROMOTING FINANCIAL AWARENESS AND STABILITY

H2020 – 687895

PROFIT KNOWLEDGE GRAPH

WORK PACKAGE NO.	**WP2**	WORK PACKAGE TITLE	**LINKED DATA LIFE CYCLE**
TASK NO.	**T2.2**	TASK TITLE	**SEMANTIC DATA MODELING, LINKING AND ENRICHMENT**
MILESTONE NO.		**5**	
ORGANIZATION NAME OF LEAD CONTRACTOR FOR THIS DELIVERABLE		**SWC**	
EDITOR		**ARTEM REVENKO (SWC)**	
CONTRIBUTORS		**IOANNIS PRAGIDIS, Eirini Karapistoli, Antonios Sarigiannidis, Panagiotis Tsintzos (DUTH), CHRISTOFOROS BOUZANIS (UoG), ANNA SATSIOU (CERTH), PETER HANECAK (EEA), CHRISTIAN BLASCHKE (SWC)**	
REVIEWERS		**CHRISTOFOROS BOUZANIS (UoG), HEIDELINDE HOBEL (SWC)**	
STATUS (F: FINAL; D: DRAFT)		**F**	
NATURE		R – REPORT	
DISSEMINATION LEVEL		PU – PUBLIC	
PROJECT START DATE AND DURATION		**JANUARY 2016, 36 MONTHS**	
DUE DATE OF DELIVERABLE: ACTUAL SUBMISSION DATE:		**DECEMBER 31, 2016** **DECEMBER 30, 2016**	

REVISION HISTORY

VERSION	DATE	MODIFIED BY	CHANGES
0.1	16-11-2016	A. REVENKO (SWC)	FIRST VERSION OF TOC
0.2	17-11-2016	A. REVENKO (SWC)	DRAFTED EXECUTIVE SUMMARY, INTRODUCTION, CONCLUSION
0.3	22-11-2016	A. REVENKO (SWC)	ADDED SECTION 2
0.4	23-11-2016	A. REVENKO (SWC)	ADDED SECTION 3
0.5	29-11-2016	A. REVENKO (SWC)	ADDED SECTION 4
0.5.1	01-12-2016	A. REVENKO (SWC)	FINALIZED EXECUTIVE SUMMARY, INTRODUCTION, CONCLUSION
0.6	05-12-2016	A. REVENKO (SWC)	FINAL OFFICIAL PRELIMINARY DRAFT VERSION OF D2.4
0.7	09-12-2016	A. REVENKO (SWC)	REVIEW FROM CHRISTOFOROS BOUZANIS RECEIVED. THE CONTENTS REVISED ACCORDINGLY.
0.7.1	12-12-2016	A. REVENKO (SWC)	REVIEW FROM HEIDELINDE HOBEL RECEIVED. THE CONTENTS REVISED ACCORDINGLY. ADDED SUBSECTION 1.1.
1.0	30-12-2016	A. REVENKO (SWC)	FINAL REVIEW, PROOFING AND QUALITY CONTROL. READY FOR SUBMISSION

List of Abbreviations

SKOS Simple Knowledge Organization System. 8, 25

SPARQL Simple Protocol and RDF (Resource Description Framework) Query Language. 10, 11, 16

STW Standard Thesaurus Wirtschaft. 5, 10–13, 16, 26

URI Unique Resource Identifier. 13, 16, 30, 31

List of Figures

List of Tables

Table of Contents

Executive Summary

The deliverable presents the results of preparing the PROFIT knowledge graph. It reports about the following results achieved in the first year of the project:

1. Creation of the PROFIT thesaurus from the two financial base thesauri;
2. Extension and usage of the PROFIT finance ontology;
3. Preparation of the first draft of the PROFIT knowledge graph;
4. Identification of the action points for future work on extending the knowledge graph.

The main contributions of the deliverable may be found in

Section 2 PROFIT thesaurus;

Section 3 PROFIT Finance ontology;

Section 4 PROFIT knowledge graph.

1 Introduction

1.1 Main Notions

This deliverable describes a process of data and knowledge modelling and contains many notions from the technical and mathematical areas. Below we introduce some of the notions that are essential for understanding the contents of the current report.

A *thesaurus* is a set of concepts with three relations defined between concepts: broader, narrower, related. The related relation is symmetric, i.e. if concept c_1 is related to c_2 then c_2 is related to c_1. The relations broader and narrower are

1. inverse of each other, i.e. if c_1 is broader than c_2 then c_2 is narrower than c_1,
2. transitive, i.e. if c_1 is broader than c_2 and c_2 is broader than c_3 than c_1 is broader than c_3.

The *top concepts* are the concepts that do not have any broaders. Leaves are the concepts that do not have any narrowers. We also using the notion of the *concept scheme* in order to be able to define hierarchies of different "nature" in a single thesaurus. Formally the concept schemes are not different from top concepts.

An *ontology* as we use it here consists of three sets: classes, relations, attributes. Later the individual concepts from a thesaurus may be assigned to classes from an ontology. We may define a class A to be a subclass of a different class B. This would imply that any individual i belonging to class A also belongs to class B. Relations define possible "links" between individuals of different classes. The relations may be symmetric or inverse. Also we may define a relation to be functional, i.e. an individual may have only one another individual that is in the specified relation to it. Attributes of individuals "link" the individuals to the values of fixed types, e.g. dates, integers, url.

For storing the thesauri we use the SKOS ontology. In addition to the three described relation on the concepts of the thesaurus the SKOS defines some attributes as, for example, primary and alternative labels of the concepts in different languages. This way SKOS thesauri are appropriate for storing multi-lingual thesauri. Moreover, SKOS thesauri use *matching* links to other thesauri to be able to define links between different thesauri.

A *knowledge graph* as we use it here is an application of an ontology to a set of individuals, assigning individuals to classes, defining relations between individuals, defining attributes of individuals. We typically use concepts from a thesaurus as individuals.

1.2 Scope of the Document

The document describes the output of the preparation of the PROFIT knowledge graph. The two necessary elements for producing the knowledge graph are a thesaurus and an ontology. Therefore, the work on preparing thesauri and ontologies was continued

in the course of preparation of this deliverable. Moreover, the process of extending controlled vocabularies is a life long process, hence the work on extending the mentioned vocabularies will be continued in the future stages of the project.

The document features a detailed description of the process of creating the thesaurus, including the steps taken to resolve the conflicts in data and a description of the scope of the thesaurus. The document features a detailed description of the ontology including the scope notes on all classes and a description of relation and attributes. Finally, the document features a description of the knowledge graph.

1.3 Relation to PROFIT Project

The controlled vocabularies play a crucial role for data processing in the PROFIT platform. Namely, such tasks as data access, search, topic analysis, sentiment analysis, recommendations, preference extraction will all rely on the concept extraction tool, and the concepts are taken from thesauri. Therefore, preparation of a detailed thesaurus covering all the related fields is essential for advancing in the project.

Ontologies offer further possibilities of enriching the data model with custom classes, relations, and attributes. However, the creation of an ontology is a manual process that affords high efforts from different domain experts[1]. In the discussion with the consortium the experts from the financial and economical domains have pointed out multiple times that the financial field is especially difficult for formal modeling and even experts in the field are not always able to reach a common conclusion. Despite these difficulties we were able to come up with a PROFIT Finance ontology. We had to accept some constraints to make this ontology real and decided to focus only on the essential part of the most general classes and relations. This approach allows for describing a broad range of entities and avoiding the deeper domain specific details. Being able to describe a broad range of entities is especially important in light of the fact that the educational materials are not ready yet and only the preliminary financial articles are ready at the moment.

The PROFIT knowledge graph is the application of the ontology to the available data instances. In our case the data instances are the concepts in the thesauri. Extension of thesauri with ontologies provides finer semantic description of the different data instances. The usage of the extended data is at least two-fold:

- the extended data is a source of information and can be output to the users of the platform;
- the finer semantics improves the performance of the different data processing tasks.

However, knowledge graph cannot be created automatically and the extended data processing possibilities can be only performed on the data instances annotated by hand.

[1] In frames of Deliverable 2.1 the experts from the consortium have clearly expressed their meaning against using such well established ontologies like FIBO in our project since the scope of FIBO is not suitable for PROFIT.

As already noted, creation of the knowledge graph involves efforts from the experts in the field. In the discussion with the consortium it was decided to create a portion of the knowledge graph revised by experts and offer the further extension of the knowledge graph to the users. This way the users may apply the knowledge obtained from the educational materials and improve their reputation in the platform.

1.4 Notations and General Notes

- In the report the text of actually executed queries is presented. The data is stored in RDF format and the queries are written in SPARQL.
- In the text the semicolon notation is used, i.e. stw:13032-1 stands for the concept identifier 13032-1 in the STW thesaurus.
- The thesauri are compatible with the SKOS[2] standard. We use "prefLabel" for the SKOS preferred label and "altLabel" for the SKOS alternative label.
- In case no ambiguity is possible we represent the concept from the thesaurus by their preferred label.

All the operations over ontologies, thesauri, knowledge graph were carried out using PoolParty [3].

[2] www.w3.org/2004/02/skos/
[3] www.poolparty.biz

Table 1: Extended STW Economics statistics

Number of concepts	6832
Number of broader/narrower relations	16175
Number of related relations	21132

Table 2: EuroVoc statistics

Number of concepts	6924
Number of broader/narrower relations	7352
Number of related relations	9628

2 PROFIT Thesaurus

2.1 Introduction

Creation of the PROFIT thesaurus (or thesauri) is one of the main goals of Task 2.2 and the whole WP 2. The thesaurus is a basis of the knowledge graph and sets a ground for annotating the textual resources, hence further processing of textual data (sentiment analysis, topic and trend analysis, recommendations, etc.). The creation of the PROFIT thesaurus started in Deliverable 2.1 submitted in Month 6. In this deliverable we worked on finalizing the thesaurus and we used it for the knowledge graph.

The partners contributed to the development of the final thesaurus by extending the thesaurus manually and providing their expertize for picking the relevant thesauri as a basis for the PROFIT thesaurus.

2.2 Underlying Thesauri

In the discussions with the partners the most relevant for the PROFIT project thesauri were identified. The base thesauri used in PROFIT are STW Economics and EuroVoc, with STW being clearly more important. DUTH have already extended STW thesaurus.

The statistics of the thesauri before merging are presented in Tables 1 and 2.

2.3 Merging Thesauri

The extended version of STW is merged with EuroVoc in order to obtain the final PROFIT thesaurus. As EuroVoc has many concepts schemes we need to create an upper concept scheme. Then we transform the existing concepts schemes into concept and add them to the new concept scheme as the top concept. The SPARQL query in Listing 1 achieves the goal.

```
PREFIX skos:<http://www.w3.org/2004/02/skos/core#>
PREFIX dcterms:<http://purl.org/dc/terms/>
delete {
    ### Delete type, dcterms:title (will become the skos:
        ↪ prefLabel), and the top concept relations (will
        ↪ become broader/narrower). ###
    ?x a skos:ConceptScheme .
    ?x dcterms:title ?title .
    ?x skos:hasTopConcept ?y .
    ?y skos:topConceptOf ?x .
}
insert {
    ### Create a concept scheme with the URI <http://vocabulary
        ↪ .semantic-web.at/CBeurovoc/EuroVoc43>. If a different
        ↪  URI should be used then change this here ###
    <http://vocabulary.semantic-web.at/CBeurovoc/EuroVoc43> a
        ↪ skos:ConceptScheme .
    <http://vocabulary.semantic-web.at/CBeurovoc/EuroVoc43>
        ↪ dcterms:title "EuroVoc 4.3"@en .
    <http://vocabulary.semantic-web.at/CBeurovoc/EuroVoc43>
        ↪ skos:hasTopConcept ?x .
    ?x skos:topConceptOf <http://vocabulary.semantic-web.at/
        ↪ CBeurovoc/EuroVoc43> .
    ### Insert data for concepts ###
    ?x a skos:Concept .
    ?x skos:prefLabel ?title .
    ?x skos:narrower ?y .
    ?y skos:broader ?x .
}
where {
    ?x a skos:ConceptScheme .
    ?x dcterms:title ?title .
    ?x skos:hasTopConcept ?y .
}
```

Listing 1: Query to shift concept schemes to concepts

At the next step, the STW thesaurus is added as a separate concepts scheme to the EuroVoc thesaurus. After that we are ready to actually fuse the two thesauri. Special attention is taken to avoid duplicate concepts. Therefore, the labels of concepts from different thesauri are compared. STW and EuroVoc clearly have overlaps, in this case the branches should not be duplicated, the branches from STW have priority.

The following queries merge concepts from EuroVoc to STW (and the other way round)

where some of their labels match. They check for skos:prefLabel and skos:altLabel matches among the concepts from EuroVoc and STW. As concepts can match multiple times the matching is done from one of the two sources to the other. This is done twice to match from EuroVoc to STW (first query) and the from STW to EuroVoc (second query). There are concepts that cannot be matched without ambiguity and they are ignored and can be found with the quality reports in PoolParty. The process in the query from Listing 2 is as follows:

- Extract labels and languages for both source thesauri separately (labels are matched language specific).
- Exclude ambiguous labels (labels that appear in more than one concept in one source).
- Unite the list from both sources and select those labels that appear in both of them (i.e. concepts can be matched via such labels). From this a list of concept matches between the two sources is created.
- Select those cases where the concepts from one source only match one concept from the other source (because we do not want to merge one concept into multiple other concepts). This is the place where the direction of matching is defined, i.e. which thesaurus is the source for matching and which is the target. This step is done in the opposite direction in the second query.
- After that all the data that needs to be deleted and transformed is extracted and the concepts are merged to their targets.
- To know afterwards which concepts were merged a dcterm:replaces triple is added with the URI of the concept that is deleted in the merging process.

```
1  PREFIX skos:<http://www.w3.org/2004/02/skos/core#>
2  PREFIX dcterms:<http://purl.org/dc/terms/>
3  delete {
4      ### Delete all data associated to the EuroVoc concepts we
           ↪ want to merge ###
5      ?EuroVoc_concept skos:prefLabel ?EuroVoc_pref .
6      ?EuroVoc_concept ?pS ?o .
7      ?s ?pO ?EuroVoc_concept .
8      ?EuroVoc_concept ?dcProp ?dcValue .
9  }
10 insert {
11     ### Add all data to the STW concepts we merge them to ###
12     ?STW_concept skos:altLabel ?EuroVoc_pref .
13     ?STW_concept ?pS ?o .
14     ?s ?pO ?STW_concept .
15     ### Add a triple so we know later which concepts were
           ↪ merged into other concepts ###
16     ?STW_concept dcterms:replaces ?EuroVoc_concept .
```

```
17 }
18 where {
19     ### Extract all EuroVoc data we need for deleting the data
       ↪ ###
20     {
21         ### prefLabel ###
22         ?EuroVoc_concept skos:prefLabel ?EuroVoc_pref
23     } union {
24         ### All other triples except concept meta-data ###
25         ?EuroVoc_concept ?pS ?o filter(?pS NOT IN (skos:
               ↪ prefLabel, dcterms:contributor, dcterms:creator,
               ↪ dcterms:created, dcterms:modified))
26     } union {
27         ### All triples where the concept appears in object
               ↪ position ###
28         ?s ?pO ?EuroVoc_concept
29     } union {
30         ### Get concept meta-data ###
31         ?EuroVoc_concept ?dcProp ?dcValue filter(?dcProp IN (
               ↪ dcterms:contributor, dcterms:creator, dcterms:
               ↪ created, dcterms:modified))
32     }
33     {
34         ### Select all those cases where a EuroVoc concept maps
               ↪  to one STW concept only (we do not want to merge
               ↪  one concept with multiple other concepts) ###
35         select (min(?STW_concept) as ?STW_concept) ?
               ↪ EuroVoc_concept
36         where {
37             {
38                 ### Select all labels that appear in both
                       ↪ EuroVoc and STW ###
39                 select distinct (min(?EuroVoc_concept) as ?
                       ↪ EuroVoc_concept) (min(?STW_concept) as ?
                       ↪ STW_concept)
40                 where {
41                     {
42                         ### Select only unique labels (i.e.
                               ↪ ignore cases where a label
                               ↪ appears in more than one concept
                               ↪ because of the ambiguity) ###
43                         select ?labels ?lang (min(?x) as ?
                               ↪ EuroVoc_concept)
44                         where {
```

```
45                                  {
46                                      ### Select from all EuroVoc
                                            ↪ concepts pref- and
                                            ↪ altLabels in lower case
                                            ↪ and language information
                                            ↪ ###
47                                      select distinct (lcase(str(?
                                            ↪ label)) as ?labels) (lang
                                            ↪ (?label) as ?lang) ?x
48                                      where {
49                                          ### We select concepts
                                                ↪ based on URI Pattern
                                                ↪ ###
50                                          ?x a skos:Concept filter(
                                                ↪ regex(str(?x),"http
                                                ↪ ://vocabulary.
                                                ↪ semantic-web.at/
                                                ↪ CBeurovoc/")) .
51                                          ?x skos:prefLabel|skos:
                                                ↪ altLabel ?label
                                                ↪ filter(lang(?label)
                                                ↪ IN ("de", "en")) .
52                                      }
53                                  }
54                              } group by ?labels ?lang having(count(?
                                    ↪ x) = 1)
55                          } union {
56                              ### The same thing for concepts from
                                    ↪ STW ###
57                              select ?labels ?lang (min(?x) as ?
                                    ↪ STW_concept)
58                              where {
59                                  {
60                                      select distinct (lcase(str(?
                                            ↪ label)) as ?labels) (lang
                                            ↪ (?label) as ?lang) ?x
61                                      where {
62                                          ?x a skos:Concept filter(
                                                ↪ regex(str(?x),"(http
                                                ↪ ://zbw.eu/stw/|http
                                                ↪ ://profit.poolparty.
                                                ↪ biz/stwen/)")) .
63                                          ?x skos:prefLabel|skos:
                                                ↪ altLabel ?label
```

```
                                                    ↪ filter(lang(?label)
                                                    ↪ IN ("de", "en")) .
64                                              }
65                                          }
66                                      } group by ?labels ?lang having(count(?
                                           ↪ x) = 1)
67                                  }
68                              } group by ?labels ?lang having(count(*) > 1)
69                          }
70                      } group by ?EuroVoc_concept having(count(?STW_concept)
                         ↪ = 1)
71              }
72      }
```

Listing 2: Query to fuse EuroVoc into STW

When we merge concepts that are directly connected to each other (either by skos:related or skos:broader/narrower) the links between them will still have the old URIs (because SPARQL operates on the each result individually and not iteratively on one after the other). This can be fixed relatively easily because we store the information about which concepts replaced which other concepts. We use this to fix the triples with the wrong URIs. The query from Listing 3 fixes the problem.

```
1  PREFIX skos:<http://www.w3.org/2004/02/skos/core#>
2  PREFIX dcterms:<http://purl.org/dc/terms/>
3  delete {
4      ?y ?p1 ?o .
5      ?s ?p2 ?y .
6  }
7  insert {
8      ?x ?p1 ?o .
9      ?s ?p2 ?x .
10 }
11 where {
12     ?x dcterms:replaces ?y .
13     {
14         ?y ?p1 ?o .
15     } union {
16         ?s ?p2 ?y filter(?p2 != dcterms:replaces) .
17     }
18 }
```

Listing 3: Query to fix mentions of old URIs

The query in Listings 4 fuses STW into EuroVoc. The query in Listing 3 may be used again to fix the old URIs.

```
PREFIX skos:<http://www.w3.org/2004/02/skos/core#>
PREFIX dcterms:<http://purl.org/dc/terms/>
delete {
    ?STW_concept skos:prefLabel ?STW_pref .
    ?STW_concept ?pS ?o .
    ?s ?pO ?STW_concept .
    ?STW_concept ?dcProp ?dcValue .
}
insert {
    ?EuroVoc_concept skos:altLabel ?STW_pref .
    ?EuroVoc_concept ?pS ?o .
    ?s ?pO ?EuroVoc_concept .
    ?EuroVoc_concept dcterms:replaces ?STW_concept .
}
where {
    {
        ?STW_concept skos:prefLabel ?STW_pref
    } union {
        ?STW_concept ?pS ?o filter(?pS NOT IN (skos:prefLabel,
            ↪ dcterms:contributor, dcterms:creator, dcterms:
            ↪ created, dcterms:modified))
    } union {
        ?s ?pO ?STW_concept
    } union {
        ?STW_concept ?dcProp ?dcValue filter(?dcProp IN (
            ↪ dcterms:contributor, dcterms:creator, dcterms:
            ↪ created, dcterms:modified))
    }
    {
        select (min(?EuroVoc_concept) as ?EuroVoc_concept) ?
            ↪ STW_concept
        where {
            {
                select distinct (min(?EuroVoc_concept) as ?
                    ↪ EuroVoc_concept) (min(?STW_concept) as ?
                    ↪ STW_concept)
                where {
                    {
                        select ?labels ?lang (min(?x) as ?
                            ↪ EuroVoc_concept)
                        where {
                            {
                                select distinct (lcase(str(?
```

```
                                    ↪ label)) as ?labels) (lang
                                    ↪ (?label) as ?lang) ?x
36                                 where {
37                                     ?x a skos:Concept filter(
                                        ↪ regex(str(?x),"http
                                        ↪ ://vocabulary.
                                        ↪ semantic-web.at/
                                        ↪ CBeurovoc/")) .
38                                     ?x skos:prefLabel|skos:
                                        ↪ altLabel ?label
                                        ↪ filter(lang(?label)
                                        ↪ IN ("de", "en")) .
39                                 }
40                             }
41                         } group by ?labels ?lang having(count(?
                            ↪ x) = 1)
42                     } union {
43                         select ?labels ?lang (min(?x) as ?
                            ↪ STW_concept)
44                         where {
45                             {
46                                 select distinct (lcase(str(?
                                    ↪ label)) as ?labels) (lang
                                    ↪ (?label) as ?lang) ?x
47                                 where {
48                                     ?x a skos:Concept filter(
                                        ↪ regex(str(?x),"(http
                                        ↪ ://zbw.eu/stw/|http
                                        ↪ ://profit.poolparty.
                                        ↪ biz/stwen/)")) .
49                                     ?x skos:prefLabel|skos:
                                        ↪ altLabel ?label
                                        ↪ filter(lang(?label)
                                        ↪ IN ("de", "en")) .
50                                 }
51                             }
52                         } group by ?labels ?lang having(count(?
                            ↪ x) = 1)
53                     }
54                 } group by ?labels ?lang having(count(*) > 1)
55             }
56         } group by ?STW_concept having(count(?EuroVoc_concept)
            ↪ = 1)
57     }
```

```
}
```

Listing 4: Query to fuse STW into EuroVoc

2.4 Resolution of Quality Issues in Fused Thesaurus

Because of the previous merging process, it can be the case that concepts now have altLabels that are identical to their prefLabels. In those cases the following query removes the altLabels. The query in Listing 5 resolves the issue.

```
PREFIX skos:<http://www.w3.org/2004/02/skos/core#>
delete { ?x skos:altLabel ?alt }
where {
    ?x skos:altLabel ?alt filter(lang(?alt) = ?lang && lcase(
       ↪ str(?alt)) = ?label) .
    ?x skos:prefLabel ?pref filter(lang(?pref) = ?lang) .
    {
        select ?x ?label ?lang
        where {
            {
                select ?x (lcase(str(?pref)) as ?label) (lang(?
                   ↪ pref) as ?lang)
                where {
                    ?x skos:prefLabel ?pref .
                }
            } union {
                select distinct ?x (lcase(str(?alt)) as ?label)
                   ↪ (lang(?alt) as ?lang)
                where {
                    ?x skos:altLabel ?alt .
                }
            }
        } group by ?x ?label ?lang having(count(*) > 1)
    }
}
```

Listing 5: Query to check label conflicts

When two concepts that are related to each other (via skos:related or skos:broader/narrower relations) and are both merged into the same target concept then there will be relations that will point back to the same concept. Those triples are removed with the query in Listing 6.

```
delete {
  ?x ?p ?x
```

```
3 }
4 where {
5   ?x ?p ?x
6 }
```

Listing 6: Query to remove loops

What is not allowed in SKOS is to have hierarchical relations between concepts (broader/narrower) and at the same time non-hierarchical ones (skos:related). For those cases the query in Listing 7 removes the skos:related triples.

```
1 PREFIX skos:<http://www.w3.org/2004/02/skos/core#>
2 delete {
3     ?y skos:related ?x .
4     ?x skos:related ?y .
5 }
6 where {
7   ?x skos:broader+ ?y .
8   ?y skos:related ?x .
9 }
```

Listing 7: Query to remove related if hierarchical

2.4.1 Manual Resolution of Label Clashes

As already mentioned there may appear a situation when a concept's prefLabel and altLabel both matches other concepts' labels. In this case there may appear an ambiguity that cannot be resolved automatically. Therefore, it was necessary to go through all the conflicts manually and either solve the conflicts or leave the concepts as is if the ambiguity is intended.

In order to fuse two concepts the query in Listing 8 is used. As follows from the names of the variables the concept `to_delete` is deleted whereas the concept `to_update` obtains all the triples from the deleted concept. For the case where the query is executed instead of describing the whole process we simply write "`to_update` consumes `to_delete`".

```
1 PREFIX skos:<http://www.w3.org/2004/02/skos/core#>
2 PREFIX dcterms:<http://purl.org/dc/terms/>
3 PREFIX eurovoc:<http://vocabulary.semantic-web.at/CBeurovoc/>
4 PREFIX stw:<http://zbw.eu/stw/descriptor/>
5 PREFIX to_delete:<http://vocabulary.semantic-web.at/CBeurovoc/{
    ↪ concept1}>
6 prefix to_update:<http://zbw.eu/stw/descriptor/{concept2}>
7 delete {
8   ### Delete all data associated to the EuroVoc concepts we
      ↪ want to merge ###
```

```
  to_delete: skos:prefLabel ?old_pref .
  to_delete: ?pS ?o .
  ?s ?pO to_delete: .
  to_delete: ?dcProp ?dcValue .
}
insert {
  to_update: ?pS ?o .
  ?s ?pO to_update: .
  ### Add a triple so we know later which concepts were merged
      ↪ into other concepts ###
  to_update: dcterms:replaces to_delete: .
}
where {
  to_delete: ?pS ?o filter(?pS NOT IN (skos:prefLabel, dcterms:
      ↪ contributor, dcterms:creator, dcterms:created, dcterms:
      ↪ modified)) filter(!isBlank(?o)).
  ?s ?pO to_delete: filter(!isBlank(?s)).
  to_delete: skos:prefLabel ?old_pref .
}
```

Listing 8: Query to_update consumes to_delete

Below we present the list of label conflicts and describe the actions taken to resolve the conflicts.

Table 3: Label Conflicts

Label	**Concepts and type of label**	**Action taken**
"Iron and steel industry" (en)	stw:13032-1 (c_1) prefLabel eurovoc:C1417 (c_2) prefLabel eurovoc:C3831 (c_3) altLabel	c_2 consumes c_1 altLabel of c_3 is removed
"Hochofen" (de)	eurovoc:C1417 (c_1) altLabel eurovoc:C3831 (c_2) altLabel	Label "Hochofen" (de) removed from c_1
"Eisen- und Stahlindustrie" (de) "Stahlerzeugung" (de) "Steel Industry" (en)	eurovoc:C1417 (c_1) altLabels stw:13029-4 (c_2) altLabels and prefLabel	c_1 consumes c_3 "Eisen- und Stahlindustrie" (de) removed from altLabel of the resulting concept since it is in prefLabel
"Provinzverwaltung" (de)	eurovoc:C5120 (c_1) altLabel stw:18808-4 (c_2) altLabel	removed from altLabels of c_1

To be continued on the next page

Table 3 Continued: Label Conflicts

"Knoblauch" (de) "Onion" (en)	stw:19024-5 (c_1) altLabels eurovoc:C1603 (c_2) altLabel and prefLabel	c_1 moved to become a narrower of c_2. Both labels removed from c_2 as c_1 is more specific
"Lauch" (de)	stw:19024-5 (c_1) altLabels eurovoc:C1604 (c_2) altLabel	Removed from altLabels of c_2
"Kohl" (de)	stw:14105-2 (c_1) altLabel eurovoc:C1604 (c_2) altLabel	"Kohl" removed from altLabels of c_1 as c_2 is more specific
"Futures market" (en)	stw:13722-4 (c_1) prefLabel eurovoc:C1787 (c_2) altLabel stw:10221-6 (c_3) altLabel	"futures market" removed from altLabels of c_2 c_1 consumes c_2
"Magrib" (et)	eurovoc:C3407 (c_1) altLabel stw:17750-0 (c_2) altLabel	"Magrib" removed from labels of c_1 since it is inconsistent, other languages do not have this altLabel, only estonian
"Mutter" (de)	eurovoc:C2239 (c_1) altLabel stw:18970-2 (c_2) altLabel	removed "Mutter" from c_1 since it is not about mutter
"Technische Dokumentation" (de)	stw:18942-0 (c_1) altLabel eurovoc:C494 (c_2) altLabel	removed "Technische Dokumentation" from c_1 since it is about information, not documentation
"Technical information" (en)	stw:18942-0 (c_1) altLabel eurovoc:C1422 (c_2) altLabel	removed "Technical information" from c_2 since c_1 is more specific
"Environmental Management" (en)	stw:18165-6 (c_1) prefLabel eurovoc:C2470 (c_2) altLabel	removed "environmental management" from c_2 since c_1 is more specific
"Inheritance" (en)	stw:16028-5 (c_1) prefLabel eurovoc:C3934 (c_2) prefLabel	c2 consumes c1
"Tropical fruit" (en)	stw:14117-2 (c_1) prefLabel eurovoc:C1120 (c_2) prefLabel	c_2 consumes c_1
"Tax" (en)	stw:11547-6 (c_1) prefLabel eurovoc:C1310 (c_2) prefLabel	c_1 consumes c_2

To be continued on the next page

Table 3 Continued: Label Conflicts

"Fascism" (en)	stw:16249-3 (c_1) prefLabel eurovoc:C971 (c_2) prefLabel	c_1 consumes c_2
"Gross national product" (en) "Bruttoinlandsprodukt" (de) "Gross domestic product" (en) "Bruttosozialprodukt" (de)	stw:10265-0 (c_1) altLabels eurovoc:C2769 (c_2) prefLabel eurovoc:C2761 (c_3) prefLabel	Removed "Gross national product", "Bruttoinlandsprodukt", "Gross domestic product", "Bruttosozialprodukt" from c_1 because others are more specific
"Handicrafts" (en)	stw:12910-1 (c_1) prefLabel eurovoc:C3619 (c_2) prefLabel stw:18927-3 (c_3) prefLabel	prefLabel removed from c_3 c_1 consumes c_2
"Rohstahl" (de) "Steel" (en)	stw:14285-2 (c_1) prefLabel eurovoc:C5065 (c_2) prefLabel	c_1 consumes c_2
"Eisen" (de)	stw:14285-2 (c_1) altLabel eurovoc:C981 (c_2) prefLabel	Removed "Eisen" from labels of c_1
"Social facilities" (en)	stw:18906-4 (c_1) prefLabel eurovoc:C3358 (c_2) prefLabel	c_1 consumes c_2
"Welfare institution" (en)	stw:18906-4 (c_1) altLabel eurovoc:C4143 (c_2) altLabel	Removed "welfare institution" from c_1
"Forage crops" (en)	stw:12939-5 (c_1) prefLabel stw:14990-2 (c_2) prefLabel	c_1 consumes c_2
"Innovation" (en)	stw:10459-1 (c_1) prefLabel eurovoc:C1439 (c_2) prefLabel	c_1 consumes c_2
"Produktentwicklung" (de)	stw:10459-1 (c_1) altLabel eurovoc:C71 (c_2) altLabel	Removed "Produktentwicklung" from c1
"Trockenfrucht" (de)	stw:14947-3 (c_1) altLabel eurovoc:C1116 (c_2) altLabel	removed altLabel "Trockenfrucht" from c_2
"Unternehmensleitung" (de) "Unternehmensfhrung" (de) "Geschftsleitung" (de)	stw:12085-6 (c_1) altLabels eurovoc:C447 (c_2) prefLabel and altLabel eurovoc:C1156 (c_3) altLabel	removed all the clashing altLabels from c_1
"Current account deficit" (en)	stw:24 (c_1) prefLabel stw:34 (c_2) prefLabel	c_1 consumes c_2
"euro zone" (en)	stw:6 (c_1) prefLabel stw:116 (c_2) altLabel	c_1 consumes c_2

To be continued on the next page

Table 4: Example of unresolved conflict

Label	**Concepts and type of label**	**Explanation**
"EMA" (et)	eurovoc:C4437 altLabel eurovoc:C2239 altLabel	In one case EMA is an abbreviation and in another case "ema" stands for "mother" (estonian)
"Betriebliche Umweltpolitik" (de)	stw:18165-6 altLabel eurovoc:C7942 altLabel	remains as is because "Betriebliche Umweltpolitik" is relevant for both concepts

Table 3 Continued: Label Conflicts

"Vocational Education" (en) "berufliche Bildung" (de)	stw:11360-1 (c_1) prefLabel and altLabel eurovoc:C795 (c_2) prefLabel eurovoc:C1074 (c_3) altLabel	c_1 consumes c_2 c_1 consumes c_3
"Distance learning" (en)	stw:11371-3 (c_1) prefLabel eurovoc:C1074 (c_2) altLabel eurovoc:C779 (c_3) prefLabel	removed altLabel from c_2 c_1 consumes c_3
"Co-determination" (en)	stw:11310-2 (c_1) prefLabel eurovoc:C6026 (c_2) prefLabel	c_1 consumes c_2
"Employee participation" (en) "Worker participation" (en)	stw:11310-2 (c_1) altLabels eurovoc:C2266 (c_2) prefLabel and altLabel	remove altLabels from c1

Despite resolving all the conflicts that appeared after merging there remain some conflicts that cannot be resolved either because of the ambiguity of words and/or abbreviations or because these conflicts are inherited from the base thesauri. In Table 4 we present a few examples of the remaining conflicts. There remain 53 label conflicts.

2.4.2 Manual Resolution of Hierarchical Cycles

Besides label conflicts there exists another type of conflict that may emerge after the merging of thesauri, namely a hierarchical conflict. We encountered two types of hierarchical conflicts:

- c_1 broader c_2 broader ... broader c_1; purely hierarchical cycle, one cannot resolve which concept is broader c_1 or c_2;
- c_1 broader c_2 broader ... broader c_3, c_1 related c_3; according to the SKOS rules it is not allowed for concepts to be in hierarchical and related relations simultaneously.

In Table 5 we present the hierarchical conflicts and actions taken to resolve the conflicts. In order to be more concise in the table the notation "$c_1 \rightarrow c_2$" stands for "c_1 broader c_2", the notation "$c_1 \leftrightarrow c_2$" stands for "c_1 broader c_2 and c_2 broader c_1", the notation "$c_1 \approx c_2$" stands for "c_1 related c_2". As prefLabel are not ambiguous anymore the concepts are represented by the prefLabel.

All the conflicts were resolved, no hierarchical conflicts left in the final PROFIT thesaurus.

Table 5: Hierarchical Conflicts

Conflicting Concepts	**Action taken**
North Africa ↔ Great Maghreb	removed Great Maghreb → North Africa
Tropical fruit → nut Tropical fruit ≈ nut	removed Tropical fruit → nut
Homes and hostels ↔ Social facilities	removed Homes and hostels → Social facilities
Occupational safety ↔ Ergonomics	removed Occupational safety → Ergonomics
tranquiliser ↔ psychotropic substance	removed psychotropic substance → tranquiliser
gift → Liability gift ≈ Liability	removed gift → Liability
Geology ↔ Earth sciences	removed Geology → Earth sciences
Fertility → Birth rate Fertility ≈ Birth rate	removed Fertility → Birth rate
Peace ↔ International commercial arbitration	removed both relations, concepts seem to not be in hierarchical relation
military equipment ↔ weapon of mass destruction	removed weapon of mass destruction → military equipment
Diplomatic mission ↔ Diplomacy	removed Diplomatic mission → Diplomacy
Occupational pension scheme ↔ pension scheme	removed Occupational pension scheme → pension scheme
transportation tariff → transport price transportation tariff ≈ transport price	removed transportation tariff → transport price
Chemical element ↔ halogen	removed Chemical element → halogen

To be continued on the next page

Table 5 Continued: Hierarchical Conflicts

Disarmament ↔ arms limitation	removed Disarmament → arms limitation
commercial bank ↔ Specialized bank	removed Specialized bank → commercial bank
Property insurance → indemnity insurance Property insurance ≈ indemnity insurance	removed Property insurance → indemnity insurance
Animal ↔ Wild animal	removed Wild animal → Animal
crime against property ↔ Delinquency	removed crime against property → Delinquency
manufactured goods ↔ Industrial product	removed manufactured goods → Industrial product
City centre ↔ town	removed City centre → town
public health → health policy → organisation of health care → public health	removed public health → health policy
Separation of powers → Legislation → Law → International Law → Public International Law → territorial law → Separation of powers	removed territorial law → Separation of powers

2.5 Result

The resulting thesaurus features two concept schemes: EuroVoc and STW. Hence one can still explore the original structures of the thesauri by using only the fused version. The top concepts contain the original 21 categories from EuroVoc and the original classification with 7 concepts from STW. In Table 7 all the covered topics are presented. All the top concepts emerge from the categorization schemes of the base thesauri. Any concepts of the thesauri may have several top concepts as broaders, i.e. belong to several categories. There are 10837 concepts and 22440 broader/narrower relation pairs, therefore there exists (22440-10837=) 11603 poly-hierarchies. The number of poly-hierarchical relations is higher than the numbers of concepts in the thesaurus.

Table 7: Top Concepts of the PROFIT thesaurus

Top Concept	**Narrower Concepts**

To be continued on the next page

Table 7 Continued: Top Concepts of the PROFIT thesaurus

POLITICS	political framework political party electoral procedure and voting parliament parliamentary proceedings politics and public safety executive power and public service
INTERNATIONAL RELATIONS	international affairs cooperation policy international balance defence
EUROPEAN COMMUNITIES	Community institutions and European civil service European Union law European construction Community finance
LAW	sources and branches of the law civil law criminal law justice organisation of the legal system international law rights and freedoms
ECONOMICS	economic policy economic growth regions and regional policy economic structure national accounts economic analysis
TRADE	trade policy tariff policy trade international trade consumption marketing distributive trades

To be continued on the next page

Table 7 Continued: Top Concepts of the PROFIT thesaurus

FINANCE	monetary relations monetary economics financial institutions and credit free movement of capital financing and investment insurance public finance and budget policy budget taxation prices
SOCIAL QUESTIONS	family migration demography and population social framework social affairs culture and religion social protection health construction and town planning
EDUCATION AND COMMUNICATIONS	education teaching organisation of teaching documentation communications information and information processing information technology and data processing
SCIENCE	natural and applied sciences humanities
BUSINESS AND COMPETITION	business organisation business classification legal form of organisations management accounting competition
EMPLOYMENT AND WORKING CONDITIONS	employment labour market organisation of work and working conditions personnel management and staff remuneration labour law and labour relations

To be continued on the next page

Table 7 Continued: Top Concepts of the PROFIT thesaurus

TRANSPORT	transport policy organisation of transport land transport maritime and inland waterway transport air and space transport
ENVIRONMENT	environmental policy natural environment deterioration of the environment
AGRICULTURE, FORESTRY AND FISHERIES	agricultural policy agricultural structures and production farming systems cultivation of agricultural land means of agricultural production agricultural activity forestry fisheries
AGRI-FOODSTUFFS	plant product animal product processed agricultural produce beverages and sugar foodstuff agri-foodstuffs food technology
PRODUCTION, TECHNOLOGY AND RESEARCH	production technology and technical regulations research and intellectual property
ENERGY	energy policy coal and mining industries oil industry electrical and nuclear industries soft energy
INDUSTRY	industrial structures and policy chemistry iron, steel and other metal industries mechanical engineering electronics and electrical engineering building and public works wood industry leather and textile industries miscellaneous industries

To be continued on the next page

Table 7 Continued: Top Concepts of the PROFIT thesaurus

GEOGRAPHY	Europe regions of EU Member States America Africa Asia and Oceania economic geography political geography overseas countries and territories
INTERNATIONAL ORGANISATIONS	Secretary-General United Nations European organisations extra-European organisations world organisations non-governmental organisations
GENERAL DESCRIPTORS	32 concepts
BUSINESS ECONOMICS	11 concepts
GEOGRAPHIC NAMES	Geographic Names Europe Asia Africa America Australia and Oceania Political and Economic Regions Peoples and Geographic Adjectives
RELATED SUBJECT AREAS	12 concepts
COMMODITIES	24 concepts
ECONOMICS	19 concepts
ECONOMIC SECTORS	35 concepts

The URIs of the concept contain the original prefix, which therefore allows to trace the origin of each concept.

2.6 Conclusion

The PROFIT thesaurus is created and ready to use in further workpages of the project. The quality assurance guarantees the absence of formal issues with the thesaurus. Even though creation of the thesaurus is life-long process and is never finished, the current state of the thesaurus is satisfactory and the first experiments (carried out in frames of D 2.2) of using (parts of) the thesaurus for document annotation shows good results.

Table 6: PROFIT thesaurus statistics

Number of concepts	10860
Number of broader/narrower relations	22464
Number of related relations	27970

There is also a disadvantage of having the original URIs of the modified concepts in the thesaurus. Namely, the user may go to the URI of a concept and the URI may resolve, however the user will only get the information about the original concept. In order to publish the actual information about the modified concept it is necessary to change the base of the URI to the actual PROFIT server and add a mapping to the original concept. This task will be performed in the second year of the project.

Table 8: PROFIT Finance ontology v 0.1 statistics

Number of classes	19
Number of relations	21
Number of attributes	2

Table 9: PROFIT Finance ontology v 0.2 statistics

Number of classes	27
Number of relations	39
Number of attributes	18

3 PROFIT Finance Ontology

3.1 Introduction

The creation of the PROFIT ontology was started in Deliverable 2.1 and continued in frames of this deliverable; one finds original description of the goals and functions of ontologies in the respective report. However, the development of an ontology is also a open-end process. Additional complexity arises from the field of interest of the PROFIT project, namely the financial field. As partners communicated in telcos, even inside experts there is no agreement regarding the ontology structure, therefore creation of an ultimate ontology is not feasible at all. Moreover, taking the limited resources into account, one can only speak of creating a very limited part of the ontology, a skeleton. This skeleton will provide a basis for the further extension of the ontology on demand, i.e. when a concrete class or relation need to be used on real data. The extension can be done either by the consortium members or suggested by the user.

3.1.1 Overview

In frames of this deliverable two partners UoG and DUTH provided their direct contributions to the ontology that helped to improve and extend the ontology. The UoG has provided an ontology-like table containing relations, attributes and 8 classes. In what follows we refer to this document as "UoG ontology". DUTH provided their contribution directly into the ontology and in discussions.

The ontology will be used to classify the main financial instances that will be found in news articles, user input (blogs, comments), and in educational materials. The key numbers are presented in Table 8 (original version) and Table 9 (current version).

Though the difference between the original version and the current version is exactly 8 classes, it is not the case that only the classes from UoG ontology were added. Actually

some classes were changed and some were removed, therefore, there are deeper difference between the different versions.

3.1.2 Structure

The diagram of the PROFIT Finance ontology is presented in Figure 1.

The classes are presented and described in Table 10. These classes constitute a high level structure for the description of the contents of financial materials. Some helper classes and relations are taken from other ontologies: Friend of a Friend, Dublin Core Metadata, FIBO Foundations, and Schema.org. The ontology is well suited to capture the main features without further details, therefore offering a broad range of usages. On the later stages of the project when the educational materials are ready the ontology could be extended in order to described more detailed differences between different materials.

Table 10: PROFIT Finance Ontology: Classes

Class	**Superclass**	**Ontology of Origin**	**Notes**
Agent		foaf	Helper class from an existing ontology to be subclassed by EconomicAgent
Bank	FinancialInstitution		Taken from UoG ontology
Company			
country	geopolitical entity	FIBO	Taken from an existing ontology
EconomicAgent	Agent		taken from the UoG ontology
EconomicFactor			taken from the UoG ontology
EconomicOrganization	Organization		suggested by partners
EconomicSector			borrowed from PROFIT thesaurus categorization
EconomicVariable			suggested by partners
FinancialInstitution	EconomicAgent		taken from the UoG ontology

To be continued on the next page

Table 10 Continued: PROFIT Finance Ontology: Classes

FinancialObject			taken from the UoG ontology
FinancialSystem			taken from the UoG ontology
geopolitical entity	physical location	FIBO	Helper class to be subclassed by country
InterestGroup	EconomicAgent		taken from the UoG ontology
InternationalFinancialInstitution	FinancialInstitution, EconomicOrganization		suggested by partners
InternationalOrganization	Organization		suggested by partners
location		FIBO	Helper class to define country
MacroEconomicFactor	EconomicFactor		taken from the UoG ontology
MicroEconomicFactor	EconomicFactor		taken from the UoG ontology
Organization	Thing	schema.org	Taken from an existing ontology to become superclass of possible organization classes
Person	Spatial Thing, Agent	foaf	To be able to represent a single person
physical location	location	FIBO	Helper class to define country
PoliticalOrganization	Organization		To be able to distinguish between different organization types

To be continued on the next page

Table 10 Continued: PROFIT Finance Ontology: Classes

Product	Thing	schema.org	To be able to classify different products of companies within different economic sectors; idea borrowed from the PROFIT thesaurus
Region	country		To classify regions of countries; the idea borrowed from the PROFIT thesaurus where many regions are contained
Spatial Thing		Geographic ontology	to be subclassed by Person
Thing		schema.org	Helper class

The whole ontology, including relations and attributes, is presented in Table 11.

Table 11: PROFIT Finance Ontology

Class	**Domain of**	**Range of**
Agent		
Bank		isRensposibleFor
Company	isRepresentedIn, produces, hasSubsidiaries, isActiveAt	isSubsidiaryOf, hasMember, isProducedBy, isEmployedBy, hasRepresentationOf
country	hasRelatedVariable, hasPart, hasProjectionOf, isMemberOf	isPartOf, participatesIn, hasProjectionIn, relatedTo

To be continued on the next page

Table 11 Continued: PROFIT Finance Ontology

EconomicAgent	hasNominalPrice, hasExpectedPriceOf, hasExpectedRate, isProducedBy, isProductOf, hasRealPrice, isEmployedBy, owns, regulates, hasAverageWage	isProducedBy, isProductOf, isPursuingPoliciesFor, isOwnedBy, isRensposibleFor, isRegulatedBy
EconomicFactor	hasCurrentValue, hasAverage, isFactorOf, hasInfluenceOn, isKeyFactorOf, hasFacevalue, isDerivedFrom, isProducedBy, isProductOf, isEpiphenomenonOf, isRegulatedBy	isFactorOf, isKeyFactorOf, isDerivedFrom, isProducedBy, isProductOf, isPursuingPoliciesFor, isEpiphenomenonOf, isRensposibleFor, regulates
EconomicOrganization		
EconomicSector	hasMember, isRegulatedBy	regulates, relatedTo, isActiveAt
EconomicVariable	relatedTo	hasInfluenceOn, hasRelatedVariable
FinancialInstitution	hasCurrentValue, hasNominalPrice, isSubsidiaryOf, hasProjectionIn, isPursuingPoliciesFor, hasRealPrice, isRensposibleFor, foundedIn, hasSubsidiaries, isRepresentedBy, hasHeadquartersIn	isSubsidiaryOf, featuresHeadquartersOf, hasProjectionOf, hasSubsidiaries, represents

To be continued on the next page

Table 11 Continued: PROFIT Finance Ontology

FinancialObject	hasCurrentValue, hasNominalPrice, hasExpectedPriceOf, hasMaximalPrice, hasExpectedRate, hasFacevalue, hasExpectedValue, isProductOf, hasValue, hasRealPrice, isOwnedBy, hasMinimalPrice	isProducedBy, owns
FinancialSystem	hasRepresentationOf, regulates	isRepresentedIn, isRegulatedBy
geopolitical entity		
InterestGroup	isProducedBy, hasAverageWage	isProductOf
InternationalFinancialInstitution		
InternationalOrganization		
location		
MacroEconomicFactor		
MicroEconomicFactor		
Organization	isSubsidiaryOf, participatesIn, hasProjectionIn, isPursuingPoliciesFor, publishes, isRensposibleFor, foundedIn, hasSubsidiaries, isRepresentedBy, hasHeadquartersIn, hasParentOrganization	isSubsidiaryOf, featuresHeadquartersOf, hasProjectionOf, isMemberOf, hasSubsidiaries, represents, hasParentOrganization

It is important to note that extending an ontology is an easy task and can be done quickly whereas changing the existing ontology is much more difficult because it may involve changing and losing the existing data and changing the software components for the representation of the knowledge graph. Therefore, the size of ontology is kept limited and the disputable elements are not added to the ontology.

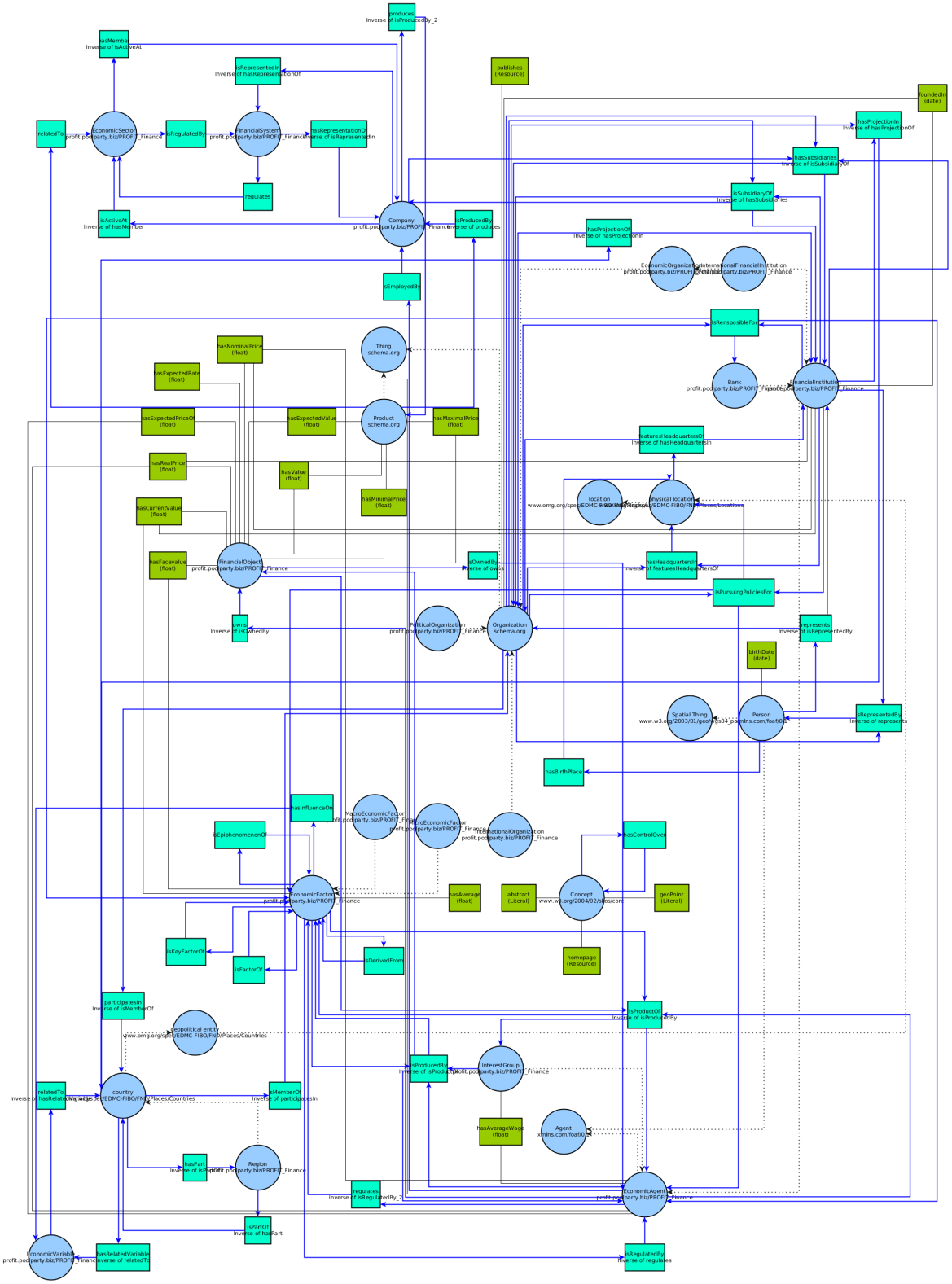

Figure 1: PROFIT Finance ontology diagram

3.2 Conclusion

The PROFIT Finance ontology is prepared and used. However, the size of the ontology is kept limited in order to provide possibilities for further extension at the later stages of the project.

4 PROFIT Knowledge Graph

4.1 Introduction

The PROFIT Knowledge Graph represents an essential selection of relevant entities, their labels (multilingual), relations and specific attributes. The data represented in the knowledge graph is highly domain specific, therefore the input from the partners, who possess the relevant expertize in the financial domain, is essentially important for creating the knowledge graph. All the data contained in the knowledge graph so far is either directly provided as input by the domain experts or is reviewed by the experts. Nevertheless, it is possible that the knowledge graph contains inaccuracies or mistakes. Correction of these mistakes is a task for the second year of the project.

Creation of the knowledge graph is a tedious task and as was agreed in the discussions between consortium members cannot be finalized within the frames of the designated resources. Therefore, the goal (in accordance with the description of work) is to create a part of the knowledge graph to demonstrate its usage and appearance in order to motivate the users of the platform to make their impact in the knowledge graph. The crowd-sourcing activities will be developed in Tasks 2.3, 3.3, 4.3 and will contribute to the knowledge graph. The goal of the consortium in the respect of knowledge graph is to provide an infrastructure to guarantee continuous sustainable development of the knowledge graph with the help of the users.

In this deliverable we focused on classifying and describing international and economic organizations and creation of the relevant part of the knowledge graph.

4.2 Overview

In Tables 12, 13, 14 parts of the knowledge graph are presented. In the first two table entities of types “Person” and “Physical location” are presented. Since these entities are not in the focus of our exercise we only present a few examples of these types. Persons are used mostly to describe the individuals representing the organizations, i.e. holding the post of a president. The locations are used to describe the location of organizations.

In Table 14 we present the whole list of classified organizations together with the information about them. Whenever the string is too long to be presented fully we use the symbol “. . . ” instead of outputting the whole string.

The knowledge graph is constantly evolving, therefore new information is constantly inflowing and by the time of the review the information presented below is likely to be extended or modified (in case any errors are found). The current information may be found at PROFIT PoolParty server.

Table 12: PROFIT Knowledge Graph: Cities

Entity	**Predicate**	**Object**
Frankfurt	featuresHeadquartersOf	European Central Bank
	abstract	...
	geoPoint	50.111805555555556 8.685944444444445
	homepage	http://www.frankfurt.de/
Washington	featuresHeadquartersOf	International Monetary Fund, World Bank, International Finance Corporation, International Economic Development Council
Istanbul	featuresHeadquartersOf	Organization of the Black Sea Economic Cooperation
Paris	featuresHeadquartersOf	Organization for Economic Cooperation and Development, Council of Europe

Table 13: PROFIT Knowledge Graph: People

Entity	**Predicate**	**Object**
Mario Draghi	hasBirthPlace	Rome
	represents	European Central Bank
	abstract	...
	birthDate	03.09.1947
Jeff Finkle	represents	International Economic Development Council

Table 14: PROFIT Knowledge Graph: Organizations

Entity	**Predicate**	**Object**
Asia–Pacific economic cooperation	class	InternationalOrganization, Thing, Organization
	abstract	...
	foundedIn	01.01.1989
	homepage	http://www.apec.org/
G8	class	InternationalOrganization, Thing, Organization
	abstract	...
European Economic and Social Committee	class	InternationalOrganization, Thing, Organization
	abstract	...
	geoPoint	50.840555555555554 4.377222222222223
	homepage	http://eesc.europa.eu/index_en.asp
European Bureau of Consumers Unions	class	InternationalOrganization, Thing, Organization
	abstract	...
Inter-American Development Bank	class	Bank, FinancialInstitution, Agent, EconomicAgent
	hasHeadquartersIn	New York
	geoPoint	38.899485 -77.030532
	homepage	http://www.iadb.org/
	abstract	...
Council of Europe	class	InternationalOrganization, Thing, Organization
	hasHeadquartersIn	Paris
	homepage	http://www.coe.int/
	foundedIn	05.05.1949
	abstract	...
United Nations	class	InternationalOrganization, Thing, Organization
	IsParentOrganizationOf	International Monetary Fund, World Bank
	hasHeadquartersIn	New York
	isRepresentedBy	Ban Ki-Moon
	homepage	http://www.un.org/en/index.html

To be continued on the next page

Table 14 Continued: PROFIT Knowledge Graph: Organizations

	foundedIn	24.10.1945
	abstract	...
European Bank for Reconstruction and Development	class	Bank, FinancialInstitution, Agent, EconomicAgent
	foundedIn	1991
	abstract	...
BusinessEurope	class	InternationalOrganization, Thing, Organization
	foundedIn	01.01.1958
	homepage	http://www.businesseurope.eu
	abstract	...
	isRepresentedBy	Jürgen R. Thumann
Belgo-Luxembourg Economic Union	class	EconomicOrganization, Thing, Organization
European League for Economic Cooperation	class	EconomicOrganization, FinancialInstitution, InternationalFinancialInstitution, Thing, Agent, Organization, EconomicAgent
	abstract	...
	foundedIn	1948
Centre for the Development of Enterprise	class	InternationalOrganization, Thing, Organization
International Standardisation Organisation	class	InternationalOrganization, Thing, Organization
	hasControlOver	Standardization
	hasHeadquartersIn	Geneva
	abstract	...
	foundedIn	23.02.1947
	homepage	http://www.iso.org/
International Trade Union Confederation	class	EconomicOrganization, Thing, Organization
	abstract	...
	foundedIn	01.11.2006
Organization for Economic Cooperation and Development	class	EconomicOrganization, Thing, Organization
	hasHeadquartersIn	Paris
	isRepresentedBy	Jose Angel Gurra

To be continued on the next page

Table 14 Continued: PROFIT Knowledge Graph: Organizations

	abstract	...
	foundedIn	16.04.1948
	homepage	http://www.oecd.org/
	publishes	https://data.oecd.org/, http://www.oecd-ilibrary.org/
European Central Bank	class	Bank, FinancialInstitution, Agent, EconomicAgent
	abstract	...
	foundedIn	01.06.1998
	geoPoint	50.1095 8.674
	homepage	http://www.ecb.europa.eu
	publishes	...
	hasHeadquartersIn	Frankfurt
	isRepresentedBy	Mario Draghi
International Monetary Fund	class	InternationalOrganization, Thing, Organization
	hasParentOrganization	United Nations
	hasHeadquartersIn	Washington
	isRepresentedBy	Christine Lagarde
	abstract	...
	foundedIn	27.12.1945
	homepage	...
	publishes	...
World Bank	class	EconomicOrganization, FinancialInstitution, InternationalFinancialInstitution, Thing, Agent, Organization, EconomicAgent
	hasParentOrganization	United Nations
	IsParentOrganizationOf	International Bank for Reconstruction and Development, International Centre for Settlement of Investment Disputes, International Development Association, International Finance Corporation, Multilateral Investment Guarantee Agency

To be continued on the next page

Table 14 Continued: PROFIT Knowledge Graph: Organizations

	hasHeadquartersIn	Washington
	isRepresentedBy	Jim Yong Kim
	abstract	...
	foundedIn	12.07.1944
	homepage	http://www.worldbank.org/
	publishes	...
International Bank for Reconstruction and Development	class	Bank, FinancialInstitution, Agent, EconomicAgent
	hasParentOrganization	World Bank
	isRepresentedBy	Jim Yong Kim
	homepage	http://www.worldbank.org/en/about/what-we-do/brief/ibrd
	abstract	...
International Centre for Settlement of Investment Disputes	class	Thing, Organization
	hasParentOrganization	World Bank
	homepage	...
	foundedIn	14.10.1966
	abstract	...
International Development Association	class	Thing, Organization
	hasParentOrganization	World Bank
	isRepresentedBy	Jim Yong Kim
	homepage	http://ida.worldbank.org/
	publishes	http://ida.worldbank.org/news
	abstract	...
International Finance Corporation	class	Thing, Organization
	hasParentOrganization	World Bank
	hasHeadquartersIn	Washington
	isRepresentedBy	Philippe Le Hourou
	abstract	...
	foundedIn	20.07.1956
	homepage	...
	publishes	...
Multilateral Investment Guarantee Agency	class	Thing, Organization
	hasParentOrganization	World Bank

To be continued on the next page

Table 14 Continued: PROFIT Knowledge Graph: Organizations

	isRepresentedBy	Keiko Honda
	abstract	...
	homepage	https://www.miga.org/
OPEC	class	Thing, Organization
	hasHeadquartersIn	Vienna
	abstract	...
	foundedIn	14.09.1960
	homepage	http://www.opec.org/opec_web/en/
	publishes	...
World Trade Organization	class	InternationalOrganization, Thing, Organization
	hasHeadquartersIn	Geneva
	isRepresentedBy	Roberto Azevdo
	abstract	...
	foundedIn	01.01.1995
	homepage	https://www.wto.org/
European Commission	class	InternationalOrganization , Thing , EU Institution , Organization
	IsParentOrganizationOf	Directorate-General for Economic and Financial Affairs, Eurostat
	hasHeadquartersIn	Brussels, Luxembourg
	isRepresentedBy	Jean-Claude Juncker
	abstract	...
	foundedIn	16.01.1958
	homepage	http://ec.europa.eu/index_en.htm
	publishes	https://ec.europa.eu/info/business-economy-euro_en
United Nations	class	InternationalOrganization , Thing , Organization
	IsParentOrganizationOf	International Monetary Fund, World Bank
	hasHeadquartersIn	New York
	isRepresentedBy	Ban Ki-Moon
	foundedIn	24.10.1945
	homepage	http://www.un.org/en/index.html

To be continued on the next page

Table 14 Continued: PROFIT Knowledge Graph: Organizations

	abstract	...
Directorate-General for Economic and Financial Affairs	class	InternationalOrganization, Thing, EU Institution, Organization
	abstract	...
	homepage	http://ec.europa.eu/economy_finance/
	publishes	...
	hasHeadquartersIn	Brussels
	hasParentOrganization	European Commission
	isRepresentedBy	Marco Buti
Eurostat	class	InternationalOrganization, Thing, EU Institution, Organization
	hasHeadquartersIn	Luxembourg
	hasParentOrganization	European Commission
	isRepresentedBy	Marianne Thyssen
	abstract	...
	homepage	http://ec.europa.eu/eurostat
	publishes	...
International Economic Development Council	class	InternationalOrganization, Thing, Organization
	hasHeadquartersIn	Washington
	isRepresentedBy	Jeff Finkle
	homepage	http://www.iedconline.org/
Organization of the Black Sea Economic Cooperation	class	EconomicOrganization, Thing, Organization
	hasHeadquartersIn	Istanbul
	foundedIn	04.04.1992
	homepage	http://www.bsec-organization.org/
	abstract	...
Paris Club	class	InternationalOrganization, Thing, Organization
	abstract	...
	homepage	http://www.clubdeparis.org

To be continued on the next page

Table 14 Continued: PROFIT Knowledge Graph: Organizations

St. Petersburg International Economic Forum	class	EconomicOrganization, Thing, Organization
	hasHeadquartersIn	St. Petersburg
	abstract	...
	foundedIn	01.01.1997
	homepage	http://forumspb.com/
UN Economic Commission for Europe	class	InternationalOrganization, Thing, Organization
	abstract	...
	foundedIn	01.01.1947
	homepage	http://www.unece.org/

4.3 Conclusion

The initial PROFIT knowledge graph is prepared and will be used to demonstrate the usage of the data as well as for the development of the related technical components of the platform. In the other tasks of this and related work packages we will develop a crowd-sourcing mechanism for extending the knowledge graph.