

# Report on the content and technical structure of the **EUPRO** Infrastructure



AIT: Austrian Institute of Technology

RISIS "Research infrastructure for research and  
innovation policy studies"

FP7, Grant agreement no: 313082

Task 6, Workpackage 6, coordinated by **AIT**  
**Austrian Institute of Technology GmbH**



# Report on the content and technical structure of EUPRO v7.4.11 (Task 6 of WP6)

Barbara Heller-Schuh\*, Michael Barber\*, Marlies Züger\* and Thomas Scherngell\*

\*AIT Austrian Institute of Technology

## Table of contents

|     |  |    |
|-----|--|----|
| 1   | Basic characteristics.....   | 3  |
| 2   | Information on substantive content of <i>EUPRO</i> .....                 | 3  |
| 2.1 | Definition and description of observations.....                          | 3  |
| 2.2 | Data acquisition and processing.....                                     | 4  |
| 2.3 | Information on all variables/indicators .....                            | 5  |
| 2.4 | Sectorial, temporal and geographical coverage.....                       | 6  |
| 2.5 | Quality and accuracy of data .....                                       | 10 |
| 3   | Legal issues encountered and access conditions .....                     | 11 |
| 4   | Technical structure of <i>EUPRO</i> .....                                | 12 |
| 4.1 | Information on the data base system .....                                | 12 |
| 4.2 | Technical variable definition.....                                       | 13 |
| 4.3 | Description of the Entity Relationship Model of <i>EUPRO</i> .....       | 14 |
| 4.4 | Interfaces for access and to other infrastructures (if applicable) ..... | 15 |
| 5   | Further planning of the opening of <i>EUPRO</i> .....                    | 15 |
| 6   | Annex .....  | 16 |

# 1 Basic characteristics

## **Name and short description of the infrastructure**

EUPRO – Database on European Framework Programmes

The EUPRO database comprises information on R&D projects and all participating organizations funded by the European Framework Programmes (FP).

## **Aim of the database (context of data acquisition)**

The EUPRO database is a significant asset of the Innovation Systems Department of AIT used for basic oriented research projects and contract research for national and international customers, such as the European Commission. It facilitates the analysis of participation patterns of organisations in different priorities of the FP and the investigation of collaborative network structures, including their evolution over time and the development of the European Research Area (ERA).

## **Legal name of operating organization**

AIT – Austrian Institute of Technology GmbH

## **Database location and type of access (access on site, online, etc.)**

EUPRO is accessible onsite at:

AIT – Austrian Institute of Technology GmbH  
Innovation Systems Department  
Donau-City-Str. 1  
1220 Vienna, Austria

# 2 Information on substantive content of *EUPRO*

## 2.1 Definition and description of observations

### **Units and definition of observations**

EUPRO covers information on:

- projects (such as project objectives and achievements, project costs, total funding, start and end date, contract type, information on the call) and
- participations (name of the participating organization, contact person with contact details, organisation type, and geographical location (NUTS2))

### **Number of observations**

EUPRO v7.4.11 comprises information on **63,860 projects** and **337,108 participations**. Table 1 disaggregates the units of observation by different FPs.

**Table 1: EUPRO database - number of projects and participations**

| <b>Programme</b> | <b>Period</b> | <b>Projects</b> | <b>Participations</b> |
|------------------|---------------|-----------------|-----------------------|
| FP1              | 1984 - 1987   | 3,283           | 7,929                 |
| FP2              | 1987 - 1991   | 3,878           | 19,068                |
| FP3              | 1990 - 1994   | 5,528           | 31,378                |
| FP4              | 1994 - 1998   | 14,562          | 67,858                |
| FP5              | 1998 - 2002   | 16,034          | 78,728                |
| FP6              | 2002 - 2006   | 9,709           | 72,012                |
| FP7*             | 2007 - 2013   | 10,866          | 60,135                |
| <b>Total</b>     |               | <b>63,860</b>   | <b>337,108</b>        |

\*Note: Until July 2012

## 2.2 Data acquisition and processing

### Where the data are retrieved from

The core data source for the construction of the EUPRO database is the CORDIS<sup>1</sup> projects database. It contains detailed information on funded projects and project participants of FPs. For the funded projects the CORDIS database lists information on project objectives and achievements, project costs, project funding, temporal location and contract type. It also provides valuable information on the project participants including the full name and address of the organisation, the participating department (partially) and the contact person.

The data on FP collaborations is publicly available. Raw data of the current version (EUPRO v7.4.11) was downloaded in July 2012. We used a wrapper – a type of web scraping program – to extract and structure the information from the website in an automated way. In order to do this, the wrapper opens each project webpage and parses the HTML content. As each webpage follows the same structure it is easy to guide the wrapper to the relevant information, which is then stored in a relational form.

### How the data are processed in terms of data cleaning

The quality of the raw data extracted from the CORDIS projects database is not generally sufficient for policy-relevant analyses. AIT has undertaken substantial efforts to improve quality and the level of standardisation of the data and to retrieve and add missing data.

Data cleaning and standardisation includes three major steps:

- identification of unique organisation name,
- identification of unique organisation type, and
- regionalisation (i.e. assignment of geographical locations to European NUTS regions).

The harmonisation of organisations names and the integration of new CORDIS data is ultimately manual, but supported by applying specific matching algorithms developed by AIT. These algorithms are based on statistical properties such as the frequency of adjacent characters in the organisation names and are used to identify similar organisation names that can be attributed to the same organisation. All algorithmically identified name matches are

---

<sup>1</sup> CORDIS (Community Research and Development Information Service, <http://cordis.europa.eu/>) is the European Commission's primary public repository and portal to disseminate information on all EU-funded research projects and their results in the broadest sense.

manually checked for accuracy. Regionalisation is based on City-to-Nuts correspondence tables.

## 2.3 Information on all variables/indicators

### Description of all variables and/or indicators that are given for the main units of observation

**Table 2: Description of variables providing information about projects**

| <b>Variable</b>  | <b>Description</b>  |
|--|---|
| RecCtrNr   | unique identifier (record control number) for each project in the database, identical with unique identifier of all projects in the CORDIS projects database  |
| Project Reference                                      | (not-unique) project index, for internal use in the European Commission (matches with Project Id in CORDA)  |
| PrgrType (code)  | code (1-7) for the names of the specific framework programme types FP1 to FP7   |
| PrgrType   | contains the corresponding names of the framework programmes (e.g., Seventh Framework Programme)  |
| Subprogramme Area                                      | full name of the subprogramme areas in each of the framework programmes   |
| PrAcr  | subprogramme acronym in the framework programmes (e.g., FP7-Health)   |
| Title  | Full title of the project   |
| ProjAcr  | (non-unique) project acronym or abbreviation of the project title   |
| Subject Index  | one or more of 68 standardized keywords; in the first three framework programmes distinct combinations of subject indices were allocated by the European Commission to projects of the same subprogram; after FP4 the allocation of subject indices to specific subprograms is more ambiguous; caution: allocation of subject indices seems sometimes arbitrary – check reliability of contents of this variable before usage |
| Other Indexes  | additional keywords to characterize project contents, freely chosen by the project team   |
| Objective, GenInfo (general information), Achievements | distinction of these fields not clearly defined, in practice any of these fields may contain objectives or an abstract of the project.  |
| Partners/Project                                       | number of participants in each project; additional generated information, calculated with the help of table “participants_FP1-7”  |
| Contract type  | different types of contracts which regulate size, financing and funding of the research projects (e.g., STREP - Specific Targeted Research Project)   |
| Start Date, End Date                                   | day, month and year of project start and end  |
| Duration (in months)                                   | duration of the project in months (data provided by cordis)   |
| Project Cost   | official project costs  |
| Project Funding  | financing contribution of the European Union; since not all projects are financed completely, figures in “Project Funding” are equal to or smaller than figures in “Project Cost”.  |
| Project Status   | current status of the project (accepted, completed, execution) given in the CORDIS projects database in July, 2012  |
| Project URL  | official website of the project   |

**Table 3: Description of variables providing information about participating organisations**

| <b>Variable</b> | <b>Description</b>   |
|-----------------|--|
| RecCtrNr        | unique identifier (record control number) for each project in the database, identical with unique identifier of all projects in the CORDIS projects database; corresponds with the entries in the field RecCtrNr in the projects table (Table 2)                                   |
| Cnr             | unique identifier (control number) assigned internally by AIT, as project participants are not uniquely indexed in the CORDIS projects database; all project-relevant information is indicated with “1”, prime contractor with “2”, and remaining participants with “3”, “4”, etc. |
| stOrg           | standardised organisation name; the EUPRO database currently covers a period of more than 30 years during which organisations have changed to mergers, acquisitions and divestitures. At the moment organisations are labeled by their most recent valid name.                     |
| stsubentity     | standardised identifier for organisational subunits; incomplete raw data on the participating subunits for FP1 to FP6; with the introduction of the Participant Identification Code (PIC) in FP7 such raw data is no longer available  |
| stOrgType       | standardised organisation type   |
| Contact Person  | name of contact person, without name affixes like titles, etc.   |
| Tel             | phone number of contact person   |
| Fax             | fax number of contact person   |
| URL             | official website of the organisation   |
| sAddress_orig   | complete address information   |
| sPostcode       | extracted postal code  |
| sCity           | extracted city   |
| sRegion         | extracted European Regions   |
| stCtry          | standardised country codes; country abbreviations are given as ISO 3166-1 Alpha-3 country codes <sup>2</sup>   |
| NUTS2           | region code according to the NUTS -2 classification scheme (revision 2010)   |
| CtryType        | type of FP association of the country in terms of eligibility for funding (FP7)  |

## 2.4 Sectorial, temporal and geographical coverage

### Information on the sectorial classifications used

**Table 4: Organisation type**

| <b>Category</b> | <b>Description</b>  |
|-----------------|---|
| IND             | industry  |
| EDU             | universities and other educational institutions   |
| ROR             | public and private research organisations   |
| GOV             | governmental institutions   |
| CON             | consultants   |
| NCL             | non-commercial/non-profit organisations   |
| OTH             | others—mainly special interest groups, like unions, chambers, inter-trade organisations, etc. |
| N/A             | not available   |

<sup>2</sup> <https://www.iso.org/obp/ui/#search/code/>

On a (no longer available) CORDIS website<sup>3</sup> subject indices were assigned to different themes. Table 5 contains the list of 68 subject indices and the themes to which they were attributed.

---

<sup>3</sup> Thematic Navigation by Subject Index, [http://cordis.europa.eu/themes/home\\_en.html#cloud](http://cordis.europa.eu/themes/home_en.html#cloud) (accessed: 13/08/2010)

**Table 5: Subject Indices and Themes**

| <b>Subject Index</b>                        | <b>Theme</b>                             |
|---|--|
| VETERINARY AND ANIMAL SCIENCES              | Agriculture and food supply              |
| AGRICULTURE                                 | Agriculture and food supply              |
| FOOD  | Agriculture and food supply              |
| AGRICULTURAL BIOTECHNOLOGY                  | Agriculture and food supply              |
| RESOURCES OF THE SEA, FISHERIES             | Agriculture and food supply              |
| MEDICINE, HEALTH                            | Biology and medicine                     |
| BIOTECHNOLOGY                               | Biology and medicine                     |
| LIFE SCIENCES                               | Biology and medicine                     |
| HEALTHCARE DELIVERY/SERVICES                | Biology and medicine                     |
| MEDICAL BIOTECHNOLOGY                       | Biology and medicine                     |
| NUCLEAR FISSION                             | Energy                                   |
| NUCLEAR FUSION                              | Energy                                   |
| FOSSIL FUELS                                | Energy                                   |
| RENEWABLE SOURCES OF ENERGY                 | Energy                                   |
| ENERGY STORAGE, ENERGY, TRANSPORT           | Energy                                   |
| ENERGY SAVING                               | Energy                                   |
| BIOFUELS                                    | Energy                                   |
| HYDROGEN AND FUEL CELLS                     | Energy                                   |
| OTHER ENERGY TOPICS                         | Energy                                   |
| CLEAN COAL TECHNOLOGIES                     | Energy                                   |
| METEOROLOGY                                 | Environment and climate                  |
| ENVIRONMENTAL PROTECTION                    | Environment and climate                  |
| RADIATION PROTECTION                        | Environment and climate                  |
| WASTE MANAGEMENT                            | Environment and climate                  |
| RADIOACTIVE WASTE                           | Environment and climate                  |
| SUSTAINABLE DEVELOPMENT                     | Environment and climate                  |
| EARTH SCIENCES                              | Environment and climate                  |
| CLIMATE CHANGE AND CARBON CYCLE RESEARCH    | Environment and climate                  |
| WATER RESOURCE MANAGEMENT                   | Environment and climate                  |
| INDUSTRIAL MANUFACTURE                      | Industry and industrial technology       |
| MATERIALS TECHNOLOGY                        | Industry and industrial technology       |
| NANOTECHNOLOGY AND NANOSCIENCES             | Industry and industrial technology       |
| INDUSTRIAL BIOTECH                          | Industry and industrial technology       |
| ELECTRONICS, MICROELECTRONICS               | Information and Communication Technology |
| INFORMATION PROCESSING, INFORMATION SYSTEMS | Information and Communication Technology |
| TELECOMMUNICATIONS                          | Information and Communication Technology |
| AUTOMATION                                  | Information and Communication Technology |
| ROBOTICS                                    | Information and Communication Technology |
| ICT APPLICATIONS                            | Information and Communication Technology |
| NETWORK TECHNOLOGIES                        | Information and Communication Technology |
| MEASUREMENT METHODS                         | Research in practice                     |
| MATHEMATICS; STATISTICS                     | Research in practice                     |
| REFERENCE MATERIALS                         | Research in practice                     |
| SCIENTIFIC RESEARCH                         | Research in practice                     |
| PROJECT MANAGEMENT METHODOLOGIES            | Research in practice                     |
| COORDINATION, COOPERATION                   | Research in practice                     |
| POLICIES                                    | Research in practice                     |



|                                 |                              |
|---------------------------------|------------------------------|
| LEGISLATION, REGULATIONS        | Research in practice         |
| FORECASTING                     | Research in practice         |
| RESEARCH ETHICS                 | Research in practice         |
| EVALUATION                      | Research outputs             |
| STANDARDS                       | Research outputs             |
| INNOVATION, TECHNOLOGY TRANSFER | Research outputs             |
| BUSINESS ASPECTS                | Research outputs             |
| INTELLECTUAL PROPERTY RIGHTS    | Research outputs             |
| SOCIAL ASPECTS                  | Social and Economic Concerns |
| EDUCATION, TRAINING             | Social and Economic Concerns |
| INFORMATION, MEDIA              | Social and Economic Concerns |
| ECONOMIC ASPECTS                | Social and Economic Concerns |
| REGIONAL DEVELOPMENT            | Social and Economic Concerns |
| EMPLOYMENT ISSUES               | Social and Economic Concerns |
| SAFETY                          | Social and Economic Concerns |
| SECURITY                        | Social and Economic Concerns |
| CONSTRUCTION TECHNOLOGY         | Transport and Construction   |
| TRANSPORT                       | Transport and Construction   |
| AEROSPACE TECHNOLOGY            | Transport and Construction   |
| SPACE AND SATELLITE RESEARCH    | Transport and Construction   |
| OTHER TECHNOLOGY                | Transport and Construction   |

### Information on the temporal coverage used

EUPRO v7.4.11 covers all EU FP projects and participants starting from 1984, with the latest update in July 2012.

### Information on the geographical coverage and classifications used

Since we have information on the geographical location of the project participants in the EUPRO database, we can analyse their geographical distribution across Europe at the country-level as well as on the regional level by assigning organisation to European NUTS regions<sup>4</sup> using NUTS classification revision 2010<sup>5</sup>.

EUPRO covers participations from the following countries:

- *EU 28 Member States*
- *Associated countries* (with science and technology cooperation agreements that involved contributing to the framework programme budget)<sup>6</sup>:  
Switzerland; Israel; Norway, Iceland and Liechtenstein; Turkey, Croatia, the Former Yugoslav Republic of Macedonia and Serbia; Albania and Montenegro; Bosnia & Herzegovina; Faroe Islands; Republic of Moldova
- *Third Countries* (countries that are not Member States, nor associated countries)<sup>7</sup>

<sup>4</sup> including the analogous territorial descriptions for Switzerland and Norway

<sup>5</sup> History of NUTS, [http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts\\_nomenclature/history\\_nuts](http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts_nomenclature/history_nuts) (accessed: 24/04/2014)

<sup>6</sup> FP7 Third Country Agreements, [http://ec.europa.eu/research/participants/data/ref/fp7/116018/fp7-third-country-agreements\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/fp7/116018/fp7-third-country-agreements_en.pdf) (accessed: 24/04/2014)

- *International Cooperation Partner Countries (ICPC)*<sup>8</sup>: Countries eligible for EU funding from Africa, Asia, Caribbean, Pacific, Eastern Europa and Central Asia (EECA), Latin America, Mediterranean Partner Countries (MPC) and Western Balkan Countries (WBC) (for the total list see Annex, Table A 1)
- *High-income countries* (normally not eligible for EU funding): United States of America, Canada, Japan, the Republic of Korea, Singapore, Australia and New Zealand, Taiwan, Hong Kong and Macao, Vatican, San Marino, Monaco and Andorra.

## 2.5 Quality and accuracy of data

### Information on the number of missing values

EUPRO v7.4.11 comprises data on 63,860 projects and 337,108 participations (see Table 1). The following two tables give an overview of the number and ratio of missing values for each variable.

**Table 6: Number and ratio of missing values of project data**

| Variable             | Missing values |       |
|----------------------|----------------|-------|
|                      | Count          | Ratio |
| RecCtrNr             | none           | 0%    |
| Project Reference    | none           | 0%    |
| PrgrType (code)      | none           | 0%    |
| PrgrType             | none           | 0%    |
| Subprogramme Area    | 15,296         | 24%   |
| PrAcr                | none           | 0%    |
| Title                | 1              | 0%    |
| ProjAcr              | 24,271         | 38%   |
| Subject Index        | 3              | 0%    |
| Other Indexes        | 54,540         | 85%   |
| Objective            | 20,592         | 32%   |
| GenInfo              | 43,383         | 68%   |
| Achievements         | 55,940         | 88%   |
| Partners/Project     | none           | 0%    |
| Contract type        | 13,759         | 22%   |
| Start Date           | 2,144          | 3%    |
| End Date             | 2,955          | 5%    |
| Duration (in months) | 8,863          | 14%   |
| Project Cost         | 27,422         | 43%   |
| Project Funding      | 22,919         | 36%   |
| Project Status       | none           | 0%    |
| Project URL          | 58,497         | 92%   |

<sup>7</sup> Cooperation with Third Country Participants in an EC funded FP7 multi-partner research project, [http://ec.europa.eu/research/participants/data/ref/fp7/90400/guideline-third-country-participants\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/fp7/90400/guideline-third-country-participants_en.pdf)

<sup>8</sup> List of International Cooperation Partner Countries (ICPC) - Annex 1 of Work Programme 2013 Cooperation, [http://ec.europa.eu/research/participants/data/ref/fp7/206006/wp-2013-annex-1-icpc-list\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/fp7/206006/wp-2013-annex-1-icpc-list_en.pdf), (accessed: 24/04/2014)

**Table 7: Number and ratio of missing values of participant's data**

| Variable       | Missing values    |       |
|----------------|-------------------|-------|
|                | Count             | Ratio |
| RecCtrNr       | 0                 | 0%    |
| Cnr            | 0                 | 0%    |
| stOrg          | 3,680             | 1%    |
| stsubentity    | 287,245           | 85%   |
| stOrgType      | 4,515             | 1%    |
| Contact Person | 68,423            | 20%   |
| Tel            | 133,451           | 40%   |
| Fax            | 142,477           | 42%   |
| URL            | 244,810           | 73%   |
| sAddress_orig  | 1,227             | 0%    |
| sPostcode      | 129,267           | 38%   |
| sCity          | 14,216            | 4%    |
| sRegion        | 54,495            | 16%   |
|                | (34,655 in EU 28) | 11%   |
| stCtry         | 1,062             | 0%    |
| NUTS2          | 13,679            | 4%    |
| CtryType       | 0                 | 0%    |

### **Estimation of data quality issues with respect to data acquisition, reliability of retrieving system**

In general, the original database CORDIS constitutes a reliable resource on all FP projects and participations. However, in some few cases information is incomplete or inconsistent, as for instance the starting and ending dates of projects, in particular for earlier FPs.

In regard to the data acquisition and retrieving system the data is sound and complete. The retrieving system used was a wrapper – a type of web scraping program.

- *Soundness*: The raw data which is extracted from the CORDIS website is copied one-on-one and simply restructured in a relational form.
- *Completeness*: The wrapper parses the HTML content of all project webpages and is guided to the relevant information. Each webpage follows a given structure. Hence, each piece of information is located in the same place and information won't be missed by the wrapper.

## **3 Legal issues encountered and access conditions**

### **Legal issues concerning access of the database**

A complete opening of the database is not foreseen, not least in order to avoid an uncontrolled proliferation of data sets and thus to ensure quality control. Precise rules for access to EUPRO will be defined as an essential precondition for the opening for EUPRO.

### **Owner of EUPRO raw data**

AIT Austrian Institute of Technology

Copyright holder for underlying EU FP raw data: © European Union, 1994-2014

Source: CORDIS, <http://cordis.europa.eu/>

### **Current practice for opening up of the database to external users**

While the database is primarily a major asset of the Innovation Systems Department of AIT, and used for both scientific research projects and contract research for national and international customers, AIT has already provided access to data in selected cases in the past, mainly for scientific purposes (joint projects and publications).

### **Legal necessities for potential opening procedures**

In general, access will be possible for joint projects and scientific publications. In addition, the provision of specific extracts of EUPRO, by way of on-site access at AIT premises, is planned. Due to the size and complexity of EUPRO, a dedicated local training and close collaboration with AIT staff is foreseen, usually requiring local presence for some 2-3 weeks.

## **4 Technical structure of *EUPRO***

### **4.1 Information on the data base system**

#### **Current data base system used**

EUPRO is realized as Microsoft Access 2007 database.

#### **Planned future technical changes concerning data base system**

None

## 4.2 Technical variable definition

### Labelling and data type of all variables

Table 8: Data type of variables providing information about projects

| Variable             | Data type |
|----------------------|-----------|
| RecCtrNr             | Number    |
| Project Reference    | Text      |
| PrgrType (code)      | Number    |
| PrgrType             | Text      |
| Subprogramme Area    | Memo      |
| PrAcr                | Text      |
| Title                | Text      |
| ProjAcr              | Text      |
| Subject Index        | Memo      |
| Other Indexes        | Memo      |
| Objective            | Memo      |
| GenInfo              | Memo      |
| Achievements         | Memo      |
| Partners/Project     | Number    |
| Contract type        | Text      |
| Start Date           | Date      |
| End Date             | Date      |
| Duration (in months) | Number    |
| Project Cost         | Number    |
| Project Funding      | Number    |
| Project Status       | Text      |
| Project URL          | Text      |

Table 9: Data type of variables providing information about participating organisations

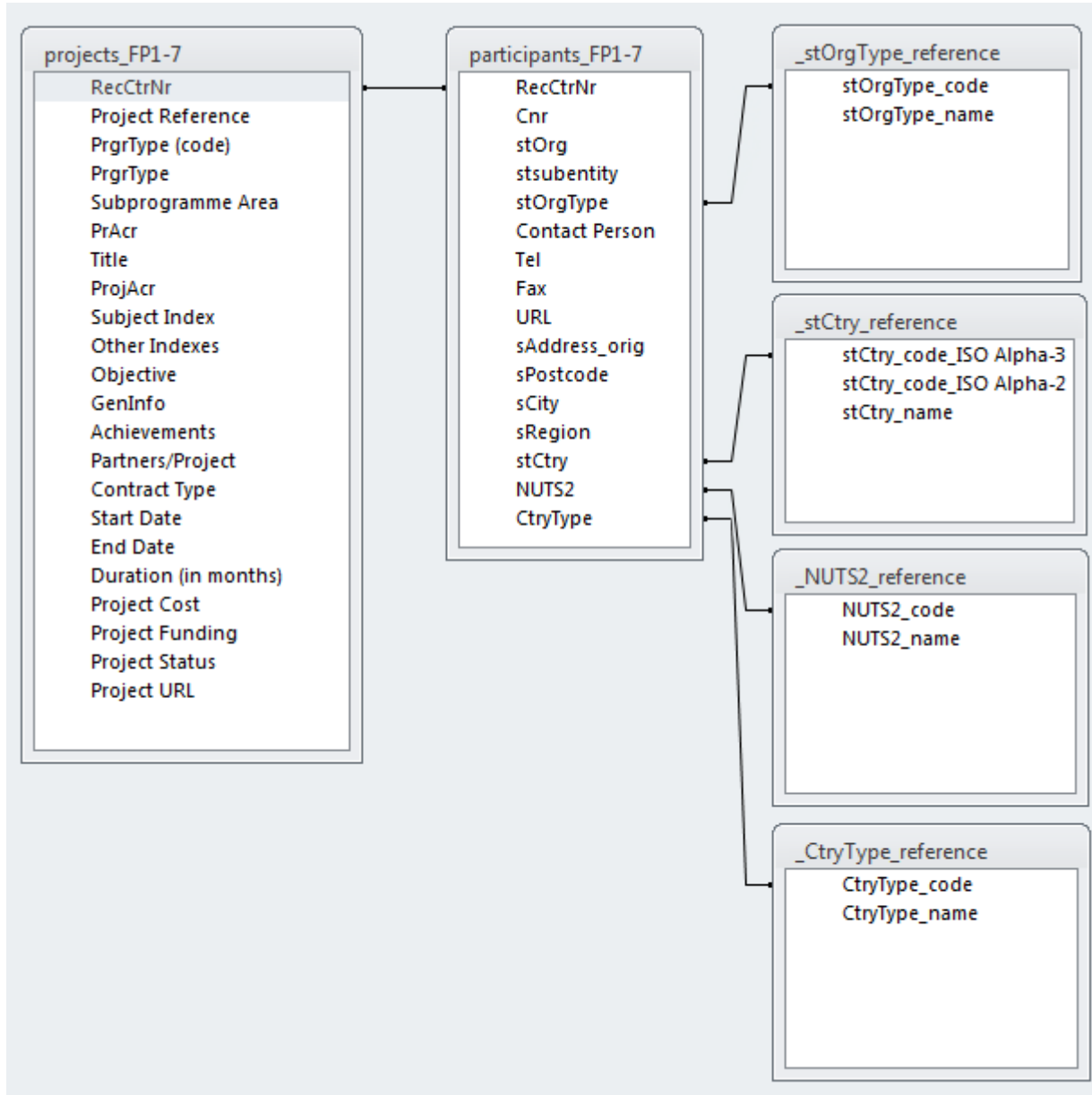
| Variable       | Data type |
|----------------|-----------|
| RecCtrNr       | Text      |
| Cnr            | Number    |
| stOrg          | Text      |
| stsubentity    | Text      |
| stOrgType      | Text      |
| Contact Person | Text      |
| Tel            | Text      |
| Fax            | Text      |
| URL            | Text      |
| sAddress_orig  | Memo      |
| sPostcode      | Text      |
| sCity          | Text      |
| sRegion        | Text      |
| stCtry         | Text      |
| NUTS2          | Text      |
| CtryType       | Text      |

### 4.3 Description of the Entity Relationship Model of EUPRO

#### Definition of single tables

EUPRO currently consists of six tables (Figure 1): *projects\_FP1-7* and *participants\_FP1-7*

Figure 1: EUPRO Entity Relationship Model



#### Relation between the tables via unique identifiers

*projects\_FP1-7* and *participants\_FP1-7* are linked via RecCtrNr (record control number), which is a unique identifier for each project in the database and identical to the unique identifier for the project in the CORDIS projects database. The four remaining tables explain abbreviations used in the respective variables.

## 4.4 Interfaces for access and to other infrastructures

### **Technical information on interfaces with other infrastructures (e.g. web interface for data search. etc.)**

At the current stage, no interfaces are planned. Potential web interfaces may be discussed and implemented during the course of the project, for instance for the provision of specific extracts of aggregated data (e.g., number of organisations/participations per country).

## 5 Further planning of the opening of *EUPRO*

### **Document concrete steps towards opening of the respective dataset**

By establishing a link to various other data sources (e.g. patent data, publications, structural data on firms, universities and PROs) the scientific potential of EUPRO shall be further enhanced. The opening date within RISIS will take place in accordance with other RISIS facilities. Accreditation and access will be formalised in accordance to Task 4 of RISIS WP6.

- New identifier at organisational level (in accordance with WP8)
- Integration of new variables (e.g. NABS code, etc.)
- Adaption to classification schemes defined in other workpackages (e.g., NUTS, organisation type, etc.)

### **Necessary updates and/or technical changes**

The EUPRO database is regularly updated by AIT, usually every second year.

In view of the growing differentiation of frameworks for R&D collaboration in Europe, not the least due to the reinforcement of multilateral arrangements among Member States, further efforts in terms of data retrieval, standardization and refinement will need to be made in the coming years in order to ensure the comprehensive coverage of EUPRO. We will thus extend the EUPRO data set by integrating additional multilateral funding data such as Joint Programming Initiatives, European Technology Platforms, Art. 185-Initiatives, EUREKA, ERANET and COST.

Potentially, a new database model will be defined in an updated version EUPRO. In this version, it is foreseen to add a new organisational table, listing organisations that are linked to projects. The new database model will be developed in accordance to other WPs, in particular the activities in WP8 on the organizational registry.

### **Changing legal conditions for accessing the dataset or parts of the dataset**

The legal access conditions will basically stay the same. Depending on the accreditation systems developed under Task 4 in WP6, changed access conditions will be defined during the course of the project.

### **Suggestions for example results to be presented at the RISIS week 2015 on the use of the dataset in order to demonstrate its value added (e.g. in conjunction with another dataset)**

It can be suggested to use a joint analysis between AIT and USI, joining data from EUMIDA/ETER with EUPRO to explain university participation in FPs by different exogenous factors. Since it is the first matching of two RISIS facilities directly contributing to address a policy relevant research questions, it may suit well to be presented at the RISIS week 2015.

## 6 Annex

**Table A 1: List of International Cooperation Partner Countries (ICPC) participating in EU FP**

| Country                    | Code | Type of Country |
|----------------------------|------|-----------------|
| ANGOLA                     | AGO  | ICPC - Africa   |
| BENIN                      | BEN  | ICPC - Africa   |
| BOTSWANA                   | BWA  | ICPC - Africa   |
| BURKINA FASO               | BFA  | ICPC - Africa   |
| BURUNDI                    | BDI  | ICPC - Africa   |
| CAMEROON                   | CMR  | ICPC - Africa   |
| CAPE VERDE                 | CPV  | ICPC - Africa   |
| CENTRAL AFRICAN REPUBLIC   | CAF  | ICPC - Africa   |
| CHAD                       | TCD  | ICPC - Africa   |
| CONGO                      | COG  | ICPC - Africa   |
| CONGO, DEMOCRATIC REPUBLIC | COD  | ICPC - Africa   |
| EQUATORIAL GUINEA          | GNQ  | ICPC - Africa   |
| ERITREA                    | ERI  | ICPC - Africa   |
| ETHIOPIA                   | ETH  | ICPC - Africa   |
| GABON                      | GAB  | ICPC - Africa   |
| GAMBIA                     | GMB  | ICPC - Africa   |
| GHANA                      | GHA  | ICPC - Africa   |
| GUINEA                     | GIN  | ICPC - Africa   |
| GUINEA-BISSAU              | GNB  | ICPC - Africa   |
| KENYA                      | KEN  | ICPC - Africa   |
| LESOTHO                    | LSO  | ICPC - Africa   |
| LIBERIA                    | LBR  | ICPC - Africa   |
| MADAGASCAR                 | MDG  | ICPC - Africa   |
| MALAWI                     | MWI  | ICPC - Africa   |
| MALI                       | MLI  | ICPC - Africa   |
| MAURITANIA                 | MRT  | ICPC - Africa   |
| MAURITIUS                  | MUS  | ICPC - Africa   |
| MOZAMBIQUE                 | MOZ  | ICPC - Africa   |
| NAMIBIA                    | NAM  | ICPC - Africa   |
| NIGER                      | NER  | ICPC - Africa   |
| NIGERIA                    | NGA  | ICPC - Africa   |
| RWANDA                     | RWA  | ICPC - Africa   |
| SAO TOME AND PRINCIPE      | STP  | ICPC - Africa   |
| SENEGAL                    | SEN  | ICPC - Africa   |
| SEYCHELLES                 | SYC  | ICPC - Africa   |
| SIERRA LEONE               | SLE  | ICPC - Africa   |
| SOMALIA                    | SOM  | ICPC - Africa   |
| SOUTH AFRICA               | ZAF  | ICPC - Africa   |
| SUDAN                      | SDN  | ICPC - Africa   |
| SWAZILAND                  | SWZ  | ICPC - Africa   |
| TANZANIA                   | TZA  | ICPC - Africa   |
| TOGO                       | TGO  | ICPC - Africa   |
| UGANDA                     | UGA  | ICPC - Africa   |
| ZAMBIA                     | ZMB  | ICPC - Africa   |
| ZIMBABWE                   | ZWE  | ICPC - Africa   |
| AFGHANISTAN                | AFG  | ICPC - Asia     |
| BANGLADESH                 | BGD  | ICPC - Asia     |
| BHUTAN                     | BTN  | ICPC - Asia     |
| CAMBODIA                   | KHM  | ICPC - Asia     |



|                                  |     |   |
|----------------------------------|-----|---|
| CHINA                            | CHN | ICPC - Asia                                   |
| INDIA                            | IND | ICPC - Asia                                   |
| INDONESIA                        | IDN | ICPC - Asia                                   |
| IRAN                             | IRN | ICPC - Asia                                   |
| IRAQ                             | IRQ | ICPC - Asia                                   |
| LAO PEOPLE'S DEMOCRATIC REPUBLIC | LAO | ICPC - Asia                                   |
| MALAYSIA                         | MYS | ICPC - Asia                                   |
| MALDIVES                         | MDV | ICPC - Asia                                   |
| MONGOLIA                         | MNG | ICPC - Asia                                   |
| MYANMAR                          | MMR | ICPC - Asia                                   |
| NEPAL                            | NPL | ICPC - Asia                                   |
| OMAN                             | OMN | ICPC - Asia                                   |
| PAKISTAN                         | PAK | ICPC - Asia                                   |
| PHILIPPINES                      | PHL | ICPC - Asia                                   |
| SRI LANKA                        | LKA | ICPC - Asia                                   |
| THAILAND                         | THA | ICPC - Asia                                   |
| VIET NAM                         | VNM | ICPC - Asia                                   |
| YEMEN                            | YEM | ICPC - Asia                                   |
| BARBADOS                         | BRB | ICPC - Carribean                              |
| CUBA                             | CUB | ICPC - Carribean                              |
| DOMINICAN REPUBLIC               | DOM | ICPC - Carribean                              |
| GUYANA                           | GUY | ICPC - Carribean                              |
| HAITI                            | HTI | ICPC - Carribean                              |
| JAMAICA                          | JAM | ICPC - Carribean                              |
| SAINT LUCIA                      | LCA | ICPC - Carribean                              |
| SURINAME                         | SUR | ICPC - Carribean                              |
| TRINIDAD AND TOBAGO              | TTO | ICPC - Carribean                              |
| ARMENIA                          | ARM | ICPC - Eastern Europe and Central Asia (EECA) |
| AZERBAIJAN                       | AZE | ICPC - Eastern Europe and Central Asia (EECA) |
| BELARUS                          | BLR | ICPC - Eastern Europe and Central Asia (EECA) |
| GEORGIA                          | GEO | ICPC - Eastern Europe and Central Asia (EECA) |
| KAZAKHSTAN                       | KAZ | ICPC - Eastern Europe and Central Asia (EECA) |
| KYRGYZSTAN                       | KGZ | ICPC - Eastern Europe and Central Asia (EECA) |
| Russian Federation               | RUS | ICPC - Eastern Europe and Central Asia (EECA) |
| TAJIKISTAN                       | TJK | ICPC - Eastern Europe and Central Asia (EECA) |
| TURKMENISTAN                     | TKM | ICPC - Eastern Europe and Central Asia (EECA) |
| UKRAINE                          | UKR | ICPC - Eastern Europe and Central Asia (EECA) |
| UZBEKISTAN                       | UZB | ICPC - Eastern Europe and Central Asia (EECA) |
| ARGENTINA                        | ARG | ICPC - Latin America                          |
| BOLIVIA, PLURINATIONAL STATE OF  | BOL | ICPC - Latin America                          |
| BRAZIL                           | BRA | ICPC - Latin America                          |
| CHILE                            | CHL | ICPC - Latin America                          |
| COLOMBIA                         | COL | ICPC - Latin America                          |
| COSTA RICA                       | CRI | ICPC - Latin America                          |
| ECUADOR                          | ECU | ICPC - Latin America                          |
| EL SALVADOR                      | SLV | ICPC - Latin America                          |
| GUATEMALA                        | GTM | ICPC - Latin America                          |
| HONDURAS                         | HND | ICPC - Latin America                          |
| MEXICO                           | MEX | ICPC - Latin America                          |
| NICARAGUA                        | NIC | ICPC - Latin America                          |
| PANAMA                           | PAN | ICPC - Latin America                          |
| PARAGUAY                         | PRY | ICPC - Latin America                          |
| PERU                             | PER | ICPC - Latin America                          |
| URUGUAY                          | URY | ICPC - Latin America                          |

|                                   |     |  |
|-----------------------------------|-----|--|
| VENEZUELA, BOLIVARIAN REPUBLIC OF | VEN | ICPC - Latin America                         |
| ALGERIA                           | DZA | ICPC - Mediterranean Partner Countries (MPC) |
| EGYPT                             | EGY | ICPC - Mediterranean Partner Countries (MPC) |
| Jordan                            | JOR | ICPC - Mediterranean Partner Countries (MPC) |
| LEBANON                           | LBN | ICPC - Mediterranean Partner Countries (MPC) |
| Libya                             | LBY | ICPC - Mediterranean Partner Countries (MPC) |
| MOROCCO                           | MAR | ICPC - Mediterranean Partner Countries (MPC) |
| SYRIA                             | SYR | ICPC - Mediterranean Partner Countries (MPC) |
| TUNISIA                           | TUN | ICPC - Mediterranean Partner Countries (MPC) |
| FIJI                              | FJI | ICPC - Pacific                               |
| PAPUA NEW GUINEA                  | PNG | ICPC - Pacific                               |
| SAMOA                             | WSM | ICPC - Pacific                               |
| SOLOMON ISLANDS                   | SLB | ICPC - Pacific                               |
| TONGA                             | TON | ICPC - Pacific                               |
| VANUATU                           | VUT | ICPC - Pacific                               |