Real-time Labour Market Information on Skill Requirements

Setting up the infrastructure for EU system

Current state of the art and next steps

Emilio Colombo
Expert Workshop
Milan, 20-21 March 2018
Major steps of the project

1. Landscaping
   - Context, specific reports, source identification

2. Early release,
   - End of 2018, only preliminary results to get a preliminary idea of what output the project can produce (Burčiak)

3. Final version,
   - All EU countries, 2020
   - Freely available (open data and methods)
Early Release

- **Area**
  - Czech Republic
  - France
  - Germany
  - Ireland
  - Italy
  - Spain
  - United Kingdom

- **Covering about 66% of European population**

- **Content**
  - Time
  - Occupation
  - Skills
  - Area
  - Industry

- **Crawling interval**
  - 9 months (April 2018-December 2018)
Time frame

- **Apr 2018**: Early Release Start Ingestion
- **Dec 2018**: Early Release
- **Sep 2018**: Full Release Start Ingestion
- **June 2020**: Full Release
Overall Data Flow - Recap

Data Ingestion - Pre-Processing - Information Extraction - ETL - Presentation Area
Landscaping

• Describe the state of the art on the use of web vacancies for labour market analysis in different countries
• Understand the context and process of data creation
• Identify the relevant sources
• Specific reports (Eures, global portals)
Source selection in landscaping

We asked International Country Experts (ICE) to identify the most relevant sources in their countries, providing key information for each of them:

- Type of job portal
- OJV approximate volume
- Geographical scope
- Structured fields
Source selection in landscaping

530 sources from landscaping

- 33 sources not available (registration required, social network, general announces)
- 24 sources linked to EURES
- 109 transnational sources (20 distinct sources)
- 364 local sources

Over 40 millions estimated OJV (high redundancy)
Overall Data Flow - Recap

Data Ingestion → Pre-Processing → Information Extraction → ETL → Presentation Area