



VICTORY

Vacancies for ICT - Online Repository

Alexandre Mateus
Policy analyst, DG Connect, Unit F4

Context

- **ICT Practitioners:** fast-growing segment of the labour force in the EU
 - Employment in ICT grew on average 4% per year since 2004
- **Shortage of ICT skills** in the labour market
 - 38% of companies trying to recruit ICT practitioners had difficulties
 - **By 2020:** over 800k potential ICT jobs that won't materialise due to lack of skilled workers
- Current data sources on ICT vacancies:
 - Show little detail (either geographical or by job category)
 - Are difficult to compare across member states
 - Are quickly outdated
- **Digital sector by definition:**
 - Virtually all advertised jobs can be found online
 - Possibility of **data collection via automated online search**

Objectives

- Monitor vacancies for ICT practitioners published online
 - **Automated search** of vacancies published on the Internet
 - Production of **statistical data** on number and characteristics of such vacancies
 - Job sub-category (ISCO)
 - Industry sector (NACE)
 - Region (NUTS 2)
 - Skills required (ESCO)
- **Continuous** data collection (weekly/monthly data updates)
- **Automated** processing
- Online **platform for visualisation** of results

History

- **2013-2014: MOVIP**
 - Small scale pilot using France as a test case

- Results:

"It is **feasible** to collect data on ICT vacancies based on automated internet search"

- But there are **many challenges**:

"A 100% automated system without human control may lead to the collection of unreliable information"

"Search methodologies should be harmonised to ensure comparability across Member States"



Challenges in automated data collection

- **Technological**
 - Permission for scraping (robots.txt)
 - Different website structures
 - NLP (Natural Language Processing) of vacancy notices
- **Contextualising** vacancy notices
 - Job category / Industry / Place of employment / Skills required / etc.
- **Structural**
 - De-duplication
 - Assessing the lifespan of vacancy notices
- **Linguistic**
 - 24 languages (at least)
 - Mapping to common taxonomies (ISCO, ESCO, NACE)

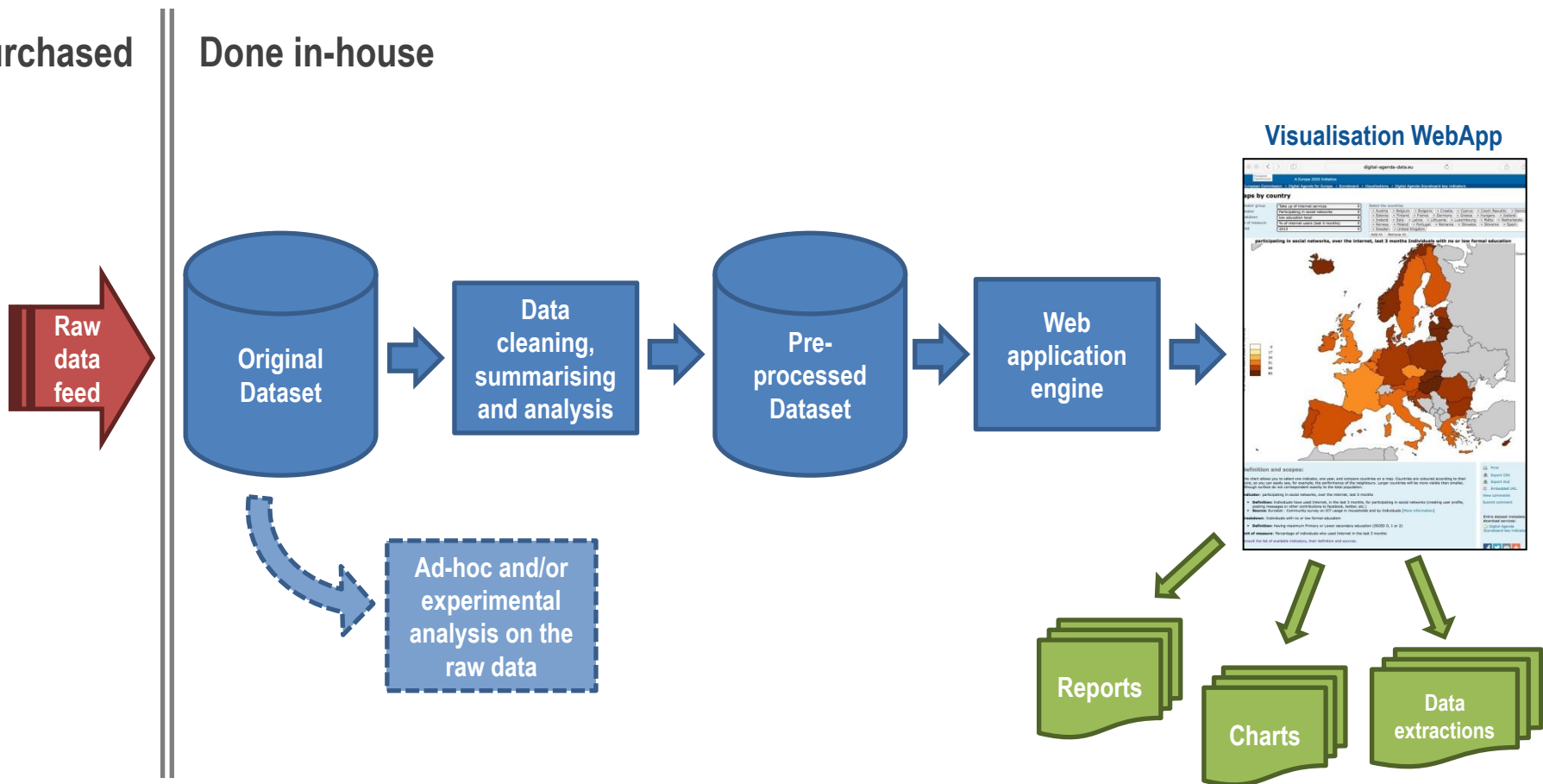
VICTORY: purchase the data

- There are companies with longstanding experience in collecting online vacancy data
 - EU coverage: around **60% of the population**
 - Plans for **expansion**
 - Crawling about **1000 job boards and websites per country**
 - **Mature technology** that deals with most mentioned challenges
- **Risk-minimising strategy**
 - **Procure** the raw data from economic operators
 - Weekly deliveries of "contextualised" raw data
 - **Develop** the infrastructure to analyse and visualise the data

Architecture

Purchased

Done in-house



Benefits

- Information for **better policymaking**:
 - Labour market policy
 - Training/education
 - Labour market mobility
- **Data and intelligence** source for:
 - Public Employment Services
 - Social partners
 - Researchers
 - Public at large
- Acquisition of experience in **big data analytics and visualisation**

Thank you

- Questions?
- Comments?