CSPA

Introduction*

* an ESS point of view

Jean-Marc Museux and Pierre Peyronnel

Eurostat
Objectives of the presentation

• Review achievements of CSPA
• Current challenges and opportunities
• Reflect and open discussion on future challenge and opportunities
Background

- Eurostat engaged in promoting the development of CSPA since the beginning as an extension of the seminal work of ESSnets CORA / CORE

- CSPA was deeply rooted in the ESS vision 2020, an ambitious programme aiming at the modernisation of the production and dissemination of EU statistics involving actively Eurostat and 28 EU NSIs

- The SERV project and the two supporting ESSnets (Sharing Common Functionalities and Implementing Shared Statistical Services) aimed to developed shared services for the ESS

- The ESS EA Reference Framework and the ESS SPRA built on CSPA (specialisation in the context of the ESS Vision 2020)
What is CSPA?

- CSPA provides a *template architecture* for official statistics, describing:
  - *What* the official statistical industry wants to achieve
  - *How* the industry can achieve this, i.e. principles that guide how statistics are produced
  - *What* the industry will have to do, adhering to the CSPA standard
Curse toward implementation

- **Conceptual**: User view, independent of physical implementation
- **Logical**: Meets requirements of CSPA Services
- **Physical**: Implemented in CSPA Services

Level of Detail → Human Oriented

Computer Oriented
CSPA Assets

CSPA Community

Collaboration and sharing

Global, open and inclusive

CSPA catalogues

Communication

Common semantic

https://www.statistical-services.org/
CSPA Service Catalogue

- Contains 35 services in different states of completion
- Landing page => opportunity for collaboration
- Recent update
  - RobotTool / NL / automated web scraping tool
  - Confidentiality Service / NZ / confidentialising microdata analysis
  - PXWeb / SE / Dissemination
  - BPM tool – GSBPM/GAMSO / IL / (to be verified)
  - Reclin / IT / Record linkage
  - Pogues / FR / Data Collection questionnaire designer
  - java-vtl / NO / VTL 1.1 engine
Services delivered by SERV

Delivered in Phase 2
Priority for Phase 3
Selected by ESSnet I3S

- Web questionnaire visualisation
- SDMX transformation
- Record Linking
- Geospatial visualisation
- SDMX RI
- Reusable Dissemination
- PX Web (OSS)

- Questionnaire generator
- SDMX registry
- Imputations
- SDMX RI

- Design outputs
- Design variables
- Design collection
- Design frame & sample
- Design processing & analysis

- Build process components
- Classify & code
- Build or enhance process components
- Build or enhance dissemination components
- Configure workflows

- Process outputs
- Classify & code outputs
- Validate outputs
- Interpret & export outputs
- Manage release of dissemination products

- Analyse error control
- Disclosure control
- Outlier detection
- Error Correction

- Disseminate dissemination products
- Disseminate Microdata access
- Manage microdata support
- Manage user support

- Services selected by ESSnet I3S
- ARC software (ADMIN data)
- PX Web (OSS)
CSPA strengths and weaknesses

Drivers

• Modernisation
• Efficiency / reuse (plug and play)
• Collaboration

Strengths

• Comprehensive
• Strong coherence (GSBPM, GSIM, CSDA)
• Consensus based
• Light governance

Weaknesses

• Few business realization
• Time to Market >> pace of change
• Lead by IT/tech roles
• Sourcing
CSPA opportunities

Drivers

• Modernisation
• Efficiency / reuse (plug and play)
• Collaboration

Opportunities

• Infrastructure as a code – infrastructure (cloud) services
• Container Technology
• Cloud Native Computing Foundation – Mesh architecture (API mediation)
• Core business logic / adaptors paradigm
ESS continued effort

Expected Outcomes of the ESSnet I3S

- Implement new services - Feed/populate the CSPA Service Catalogue
- Reality check of the guidelines with respect to implementation -> cook book
- Realize co-development and reuse

Keep fostering collaboration and the community in the ESS by:

- Continued use of the CSPA Service Catalogue as a convergence point
- Set up an ESS Expert Group on sharing tools
Future challenges

Trusted Smart Statistics

- Need for agility – adapt to the change in data ecosystem
- Transparency and reproducibility
- Technology disrupting the data gathering phase
- New methodological paradigm
- Interdisciplinary
Hourglass model

INTERMEDIATE DATA & META-DATA

Low-level processing by technology-specific experts

High-level processing methodology to be developed by domain-specific expert statisticians

Interests S-Layer

Convergence C-Layer

Data D-Layer

Heterogeneity, Complexity, Multiplicity, Variability of data sources across different data providers

parsimony, stability few common definitions

Heterogeneity, Complexity, Multiplicity, Variability of statistical indicators across different SO
Current trends

Distributed computing

• Pushing computation out (at the edge)
• Security and privacy preserving paradigm
• Micro services and CI/CD

Development of shared platform

• ESS web intelligence HUB
• UN Global Platform
Thank you