CSPA in a Nutshell: the Now

- Collect
- Process
- Analyse
- Disseminate

Canada

Sweden
CSPA in a Nutshell: the Future

Collect | Process | Analyse | Disseminate

Canada

Sweden

?
CSPA - Hot Topics from Wiesbaden 2017

- Is CSPA about IT or about Statistics?
- How can we share more?
- What is still missing? What updates do we need?
CSPA = Statistics or IT?

• CSPA is about **statistics first** and about **IT second**.

• First of all, we want to **unify & share processes and services**.

• Sharing methods might require emphasis on a **methodology architecture**.
Two compatible views on CSPA Services

What IT people see

What methodologists see

Related to one or more GSBPM sub processes

Input (specified in terms of GSIM objects)

Output (specified in terms of GSIM objects)

ARIMA, ...

Adapters (for data sources, transformation, ...)

Runnable Code

Container technology, Webservices, Applications, ...

R, Python, ...

input

output

SDMX, DDI, ...

...
CSPA = Statistics + IT!

• The sharing begins at the level of core logic that delivers value to the statistical process.

• IT people will get “any” implementation of core logic to work by creating adaptors that will make it fit for any IT environment in any NSI.

• Of course, this will be made easier by adopting some guidelines to improve CSPA Compliance.
Guidelines

Statistical

• Documentation (max 10 pages, dumbed down) in English.
• Input / output description (**GSIM**)
• Methodology (Core Logic) & position in **GSBPM**
• Use of standard models

IT

• Source Code (if distributed) in English
• Minimal technical dependencies & use of Adaptors
• Not dependent on specific data storage / internal data structures, utilizes **GSIM**
• Uses modern communication patterns
• Documentation
  • Distribution mechanism
  • License
  • Level of support
• Self-evaluation along some evaluation criteria regarding shareability:
  • Production-readiness
  • Granularity
  • Adaptability
  • Multilingualism
  • Openness (code legibility, open source, open standards...)

also known as
CSPA Compliance
Standardized Input / Output

- The key to creating a platform of reusable services benefits from a common data structure, both for input and output.
- The data structure should not be specific for a certain method, but should be described in a common and generic way.

CSPA - Hot Topics from Wiesbaden 2017

• Is CSPA about IT or about Statistics?

• How can we share more?

• What is still missing? What updates do we need?
What can we share?

• Methods
• Trained Models (Machine Learning, Auto Classification models, ...)
• Implemented methods (Core logic)
• Implemented services (Core logic + adaptors = CSPA service)
Communication

- We need a (short) brochure that shows the Goal & Business Value of CSPA
- Improve consistency on the website, remove outdated materials
- Verify if every role has documentation on the right level.

<table>
<thead>
<tr>
<th></th>
<th>Investor</th>
<th>Designer</th>
<th>Builder</th>
<th>User</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Business</strong></td>
<td>Brochure</td>
<td>Brochure</td>
<td>Brochure</td>
<td>Brochure</td>
</tr>
<tr>
<td><strong>Information</strong></td>
<td></td>
<td>Cookbook with methodology guidelines</td>
<td></td>
<td>User manual</td>
</tr>
<tr>
<td><strong>Applications</strong></td>
<td></td>
<td></td>
<td>Cookbook with IT guidelines</td>
<td>User manual</td>
</tr>
<tr>
<td><strong>Technology</strong></td>
<td></td>
<td></td>
<td>Examples</td>
<td></td>
</tr>
</tbody>
</table>
Should we update the Service Catalogue?

• Examples: Algorithmia, UN GWG Method Marketplace

• Considerations
  • User-friendliness: nice “look and feel”, easily searchable
  • Visibility: visible to anyone or to registered users?
  • Contributions: moderated or not?

• Can we align with Sharing Tools Group?
## CSPA Service Catalogue Home Page

### CSPA Service Catalogue

<table>
<thead>
<tr>
<th>Name</th>
<th>Owner</th>
<th>CSPA Compliance</th>
<th>Main GSBPM Sub-process</th>
<th>Available Implementations</th>
</tr>
</thead>
<tbody>
<tr>
<td>List Statistical Classifications</td>
<td>Statistics Norway (SSB)</td>
<td>fully</td>
<td>2.2 Design variable descriptions</td>
<td></td>
</tr>
<tr>
<td>R Engine</td>
<td>Statistics Canada</td>
<td>fully</td>
<td></td>
<td>R Engine impl</td>
</tr>
<tr>
<td>Linear Rule Checking</td>
<td>CBS (Statistics Netherlands)</td>
<td>partially</td>
<td>5.3 Review and validate</td>
<td>Linear Rule Checking_imp: R Implementation</td>
</tr>
</tbody>
</table>
Browse all algorithms:

- **Top Rated**
- **Most Called**
- **Recently Added**

- **Colorful Image Colorization**
  - deeplearning/ColorfulImageColorization
  - Colorizes given black & white images.

- **Summarizer**
  - nlp/Summarizer
  - Summarize english text
CSPA - Hot Topics from Wiesbaden 2017

- Is CSPA about IT or about Statistics?
- How can we share more?
- What is still missing? What updates do we need?
What is missing? What updates do we need?

• Links to and consistency with other standards (GAMSO, GSBPM, GSIM)

• Add references to Methodology Architecture

• Updates to the CSPA Specification include:
  • The Adaptor concept needs to be explained
  • The concept of “service” needs further explanation
  • Examples of architectural patterns

• Creation of CSPA Cookbooks that explain how to achieve a high degree of Shareability.