United Nations Economic Commission for Europe
Statistical Division

Workshop on the Modernisation of Statistical Production and Services
November 19-20, 2014

Implementing
the Common Statistical Production Architecture

Implementing CSPA in the NSI - It is a Journey…

Matjaž Jug matjaz.jug@unece.org

Agenda

- Overview of the project
- Retrospective
- Catalogue Demo
- Questions
- In the Spotlight
- Findings/Recommendations
- Next Steps
- General discussion

Project Objectives

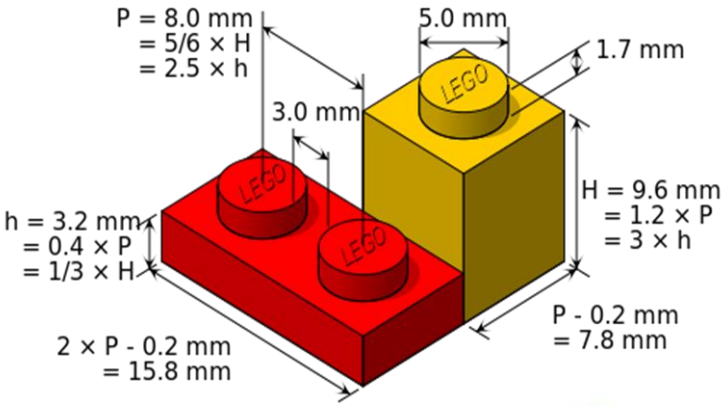
- To implement CSPA in practice by creating CSPA-compliant services that can be shared between processes and organisations (including resolving any specific licensing issues)
- To develop the resources necessary to support CSPA implementation, including training materials, and the proposed catalogue of services and other artefacts
- To further test the applicability of the GSIM, and, if necessary, to suggest further refinements to that model for a possible future revision.

Outputs

- 2013
  - CSPA v1.0
  - Proof of Concept
    - Paper
    - Videos
  - Example services
- 2014
  - CSPA v1.1 (not 2.0!)
  - 8 Implemented Services (5 implemented & reviewed and 3 in progress)
  - Supporting Tools
    - Global Service Catalogue
    - Technical Repository
    - Guidelines, Training materials, SW legal framework, agreed process with ModCom Standards for GSIM standards

Supporting Teams

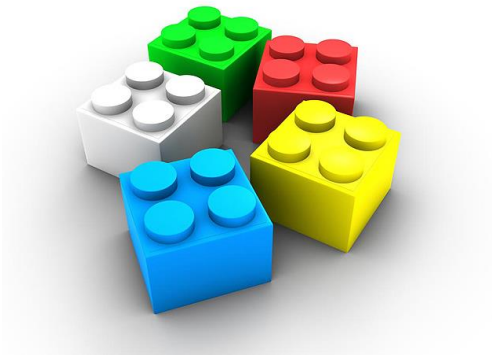
- Architecture Working Group:
  Australia, Austria, Canada, France, Italy, Mexico, Netherlands, New Zealand, Sweden, Turkey, Eurostat



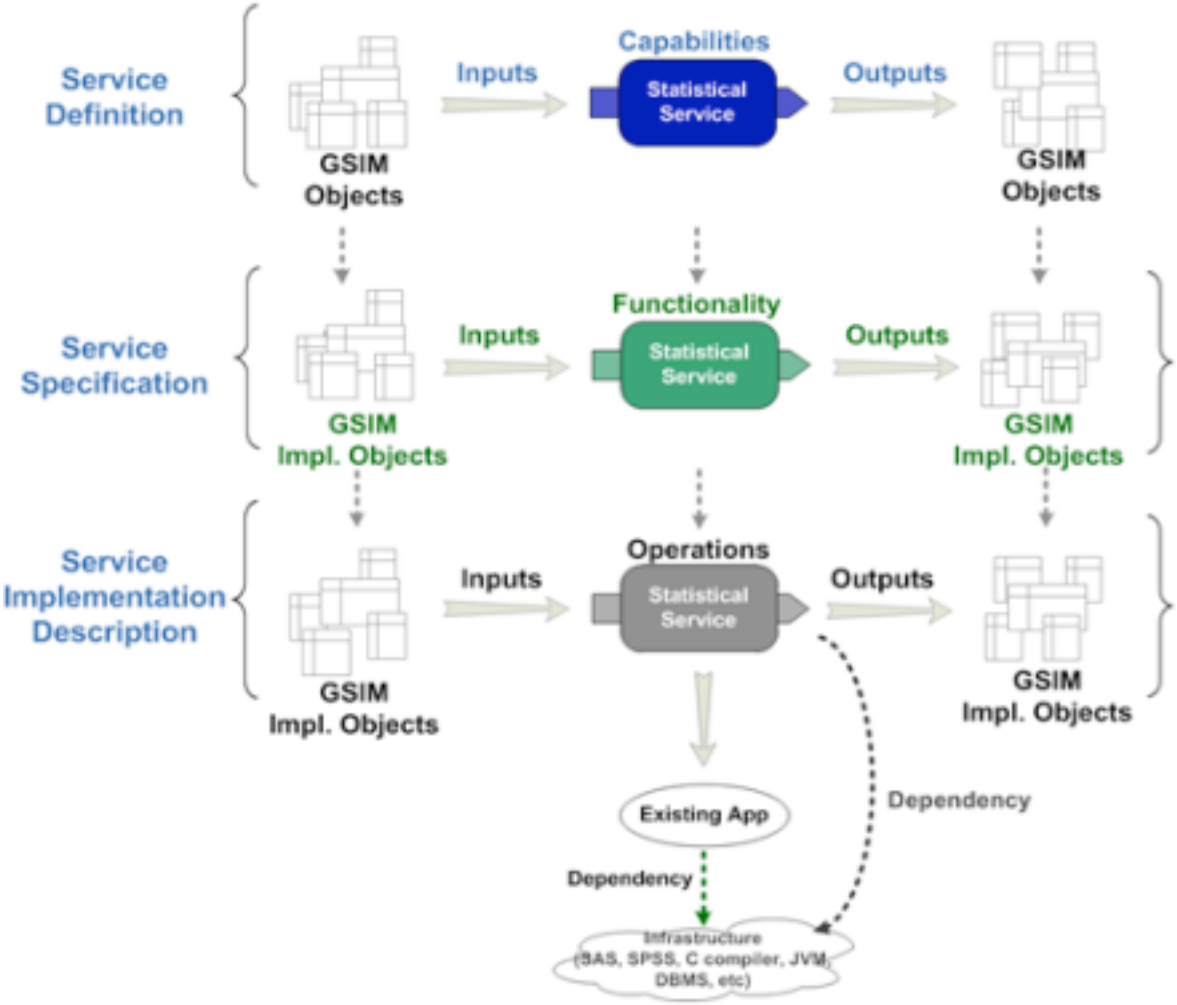
- Catalogue team:
  Australia, Canada, Italy, Hungary, New Zealand, Netherlands, Romania, Turkey, Eurostat

Services Implementation Teams

- Seasonal Adjustment – France, Australia, New Zealand
- Confidentiality on the fly – Canada, Australia
- Error correction – Italy
- SVG Generator – OECD
- SDMX transform – OECD
- Selecting sample from business register – Netherlands
- Editing components – Netherlands
- Classification Editor – Norway



Approach


Service Definition
GSIM Objects
Inputs
Capabilities
Statistical Service
Outputs
GSIM Objects
Service Specification
GSIM Impl. Objects
Inputs
Functionality
Statistical Service
Outputs
GSIM Impl. Objects
Service Implementation Description
GSIM Impl. Objects
Inputs
Operations
Statistical Service
Outputs
GSIM Impl. Objects
Existing App
Dependency
Dependency
Infrastructure (SAS, SPSS, C compiler, JVM, DBMS, etc)

Progress - Service Definition

Deliverables	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov
Service Definition Sprint #1 - #3	Create	Create	Create	Complete					Review		
Confidentialization Analysis of Microdata	X	X	X	X	X	√			X		
Error Correction	X	X	X	√					X		
Linear Error Localisation	X	X	X	√					X		
Linear Rule Checking	X	X	X	√					X		
List Statistical Classifications	X	X	X	√					X		
Retrieve Statistical Classifications	X	X	X	√					X		
SDMX Transform (Dataset Transformation)	X	X	X	X	√				X		
SVG Generator (Publication Graph Generating Engine)	X	X	X	X	√				X		
Sample Selection	X	X	X	√					X		
Seasonal Adjustment	X	X	X	√					X		

Definition phase changed original number of services from 8 -> 10

Progress - Service Specification

Deliverables	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov
Sprint #2 - **Service Specification** Sprint #4 - #6			Create	Create	Complete				Review		
Confidentialization Analysis of Microdata			X	X	X	X	X	X	X	√	
Error Correction			X	X	√				X		
Linear Error Localisation			X	X	√				X		
Linear Rule Checking			X	X	√				X		
List Statistical Classifications			X	X	X	√			X		
Retrieve Statistical Classifications			X	X	X	√			X		
SDMX Transform (Dataset Transformation)			X	X	X	X	X	√	X		
SVG Generator (Publication Graph Generating Engine)			X	X	X	X	√		X		
Sample Selection			X	X	√				X		
Seasonal Adjustment			X	X	X	X	√		X		

Specification phase challenging for Confidentialisation Analysis Service(s)

Progress - Service Implementation

Deliverables	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov
Service Implementation Sprint #7 - #15						Create	Create	Create	Complete		
Confidentialization Analysis of Microdata						X	X	X	X	X	
Error Correction						X	X	X	√		
Linear Error Localisation						X	X	X	√		
Linear Rule Checking						X	X	X	√		
List Statistical Classifications						X	X	X	stopped		
Retrieve Statistical Classifications						X	X	X	stopped		
SDMX Transform (Dataset Transformation)						X	X	X	X	X	
SVG Generator (Publication Graph Generating Engine)						X	X	X	X	√	
Sample Selection						X	X	X	X	√	
Seasonal Adjustment						X	X	X	X	X	

Classification services implementation stopped as internal project didn't get approval.

Architecture Working Group

- Meetings every fortnight.
- 25 Definition, Specification and Implementation reviews.
- 30 architectural and implementation issues.

- Update of CSPA framework.

Retrospective

Liked, Lacked, Learned, Longed for

http://www1.unece.org/stat/platform/display/pandp/Retrospective+for+CSPA+Implementation

LIKED

- The project was properly rooted in all management levels, assuring support and implementation in the NSI's that are participating.
- The actual building of a number of services helped us identify the things that still needed solving.
- The insight that this project was about high level models and methodology as well as low level IT and interoperability created feasible outcomes.
- The use of the wiki and webex made it cheap, easy and efficient to contribute to discussions and there was always someone willing to respond and comment on queries and observations.
- Templates made it easy to get started and to compare across the services
- There was significant commitment to making this work, with people accepting that not everything could be resolved now, a lot has been learned, and this has been captured.
- Really starting to identify the implementation support required, what should help resolve the issues identified

LACKED

- Governance underestimated, both central and local.
- Besides GSBPM, we needed a more detailed shared business service / capability architecture to provide a clearer context for the CSPA services.
- True collaborative development of services, where multiple NSOs can contribute as designers\builders rather than as currently where any one role is assumed entirely by a given NSO.
- There was not always a good match between commitments made at the executive level with regard to services and resources required to make a successful attempt at building the required services.
- We failed to define the project organization and governance in a sufficiently precise way, so the roles were not that clear.
- Internal administration processes change when reuse becomes the norm.

LEARNED

- With CSPA there is a really large organisational change process associated.
- We should have started from the beginning to think about the physical implementation, we are late now.
- Tension between CSPA agnostic approach to implementation and desire to be implementation specific.
- We need to create an environment where service builders are willing to contribute.
- This is a good way to describe services that we can also use in our own organisation.
- Working virtually in a team with members in different times zones caused difficulties. Using e-mails, issues that would have been solved within minutes if working on-site together took days to solve.
- The lack of a canonical GSIM implementation will continue to undermine efforts to get true interoperability of services.
- WebEx is not a project management tool.

LONGED FOR

- A interface working group that would have specified the working parts like the consortium that specified the technical stuff of the mobile GSM system.
- Alignment on NSI project portfolios. Ideally in November but April will do (we could use the new MSIS meeting, the WMSP for that)
- A high level description on how these services were expected to be "wired" together.
- Scheduling 1-2 'sprint weeks' would greatly improve the project outcomes. Having one at the start and one halfway along would help introduce everyone involved, sort out roles and responsibilities and assess progress being made.
- Top-down design of CSPA services to be implemented, i.e. the development of concrete services implementing agreed methods that are used by NSIs…
- The catalogue is an important piece of the CSPA picture. However, there should be a plan for populating it.
- More countries participating.

CSPA
Global Artefact Catalogue

(deliverable in conjunction with ModCom Production and Methods)

Jenny Linnerud

jal@ssb.no

CSPA Global Artefact Catalogue

- A primary aim of CSPA is to support efficient sharing and reuse of process patterns, information and services at an organization and international level.
- One key requisite in achieving this goal is an ability to reliably and efficiently discover what is available for reuse to support a particular business need.
- This includes an ability to efficiently assess whether a potentially reusable artefact is, in fact, "fit for purpose" in practice when it comes to supporting that particular business need.

Five layers of the CSPA Global Artefact Catalogue

Layer	Description
Virtual Standards Help Desk	*Frameworks e.g. GSBPM, GSIM, CSPA* *International Standards e.g. DDI, SDMX*
Business Plans and Interests	*Using ideas from the Database of International Statistical Activities (DISA)*
Business Capabilities	*Work ongoing in the Statistical Network.* *Planned delivery: early 2015.*
Statistical Services	*CSPA Service Description* *CSPA Service Specification* *CSPA Service Implementation*
Technical/Supporting Services	*Technical Repository*

UNECE Virtual Standards Helpdesk

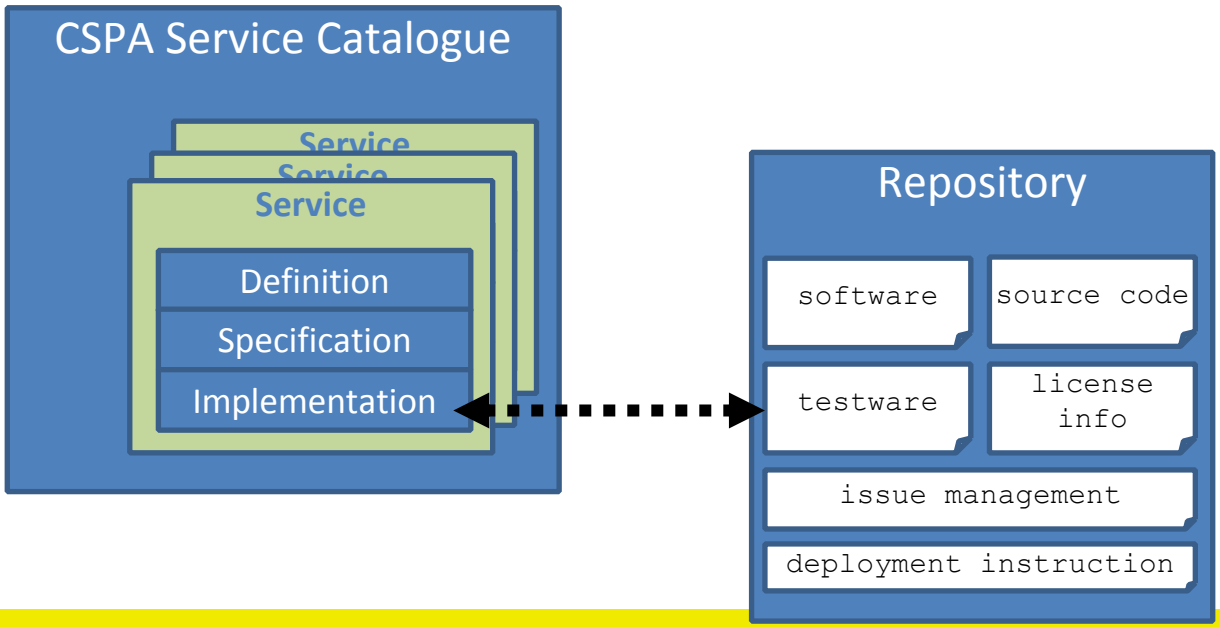
Welcome to the UNECE Virtual Standards Helpdesk

This wiki provides a "one-stop shop" for access to information about the standards necessary for the modernisation of official statistics. It is an initiative of the High-Level Group for the Modernisation of Statistical Production and Services (HLG).

The standards referenced here are cross-cutting, supporting the modernisation of all types of statistical production, and are endorsed by the HLG.

CSPA Technical Repository

- Is accessible from the CSPA service catalogue
- Contains everything necessary to deploy a service within an NSI
- Contains software and deployment instructions
- Might contain source code (depending on license).



CSPA
Statistical Service Catalogue

The CSPA Statistical Service Catalogue currently contains 5 statistical services

Demo

http://www1.unece.org/stat/platform/display/pandp/CSPA+Statistical+Service+Catalogue

In the Spotlight

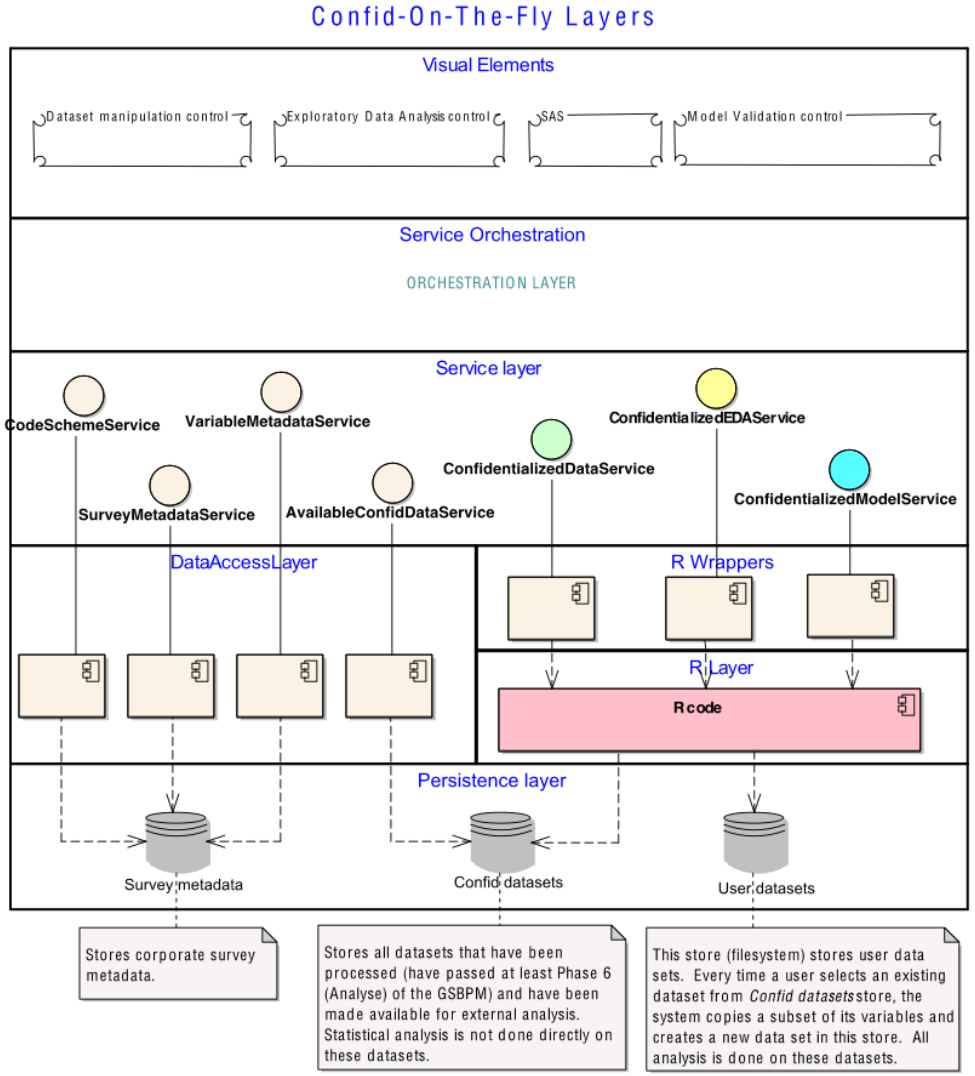
Design - Confidentiality Analysis (StatCan)

Rob McLellan

robert.McLellan@statcan.gc.ca

Confid-on-the-Fly - Schematic

- Service layer contains several methods supporting dataset and metadata locating, model generation
- Technology layer based on R, web services
- Data layer provides persistence – strong linkage to local environment(s)
- Graphical interface provided by consuming environment (Agency)
- Support for workflow orchestration depending on local environment (e.g. SOA)



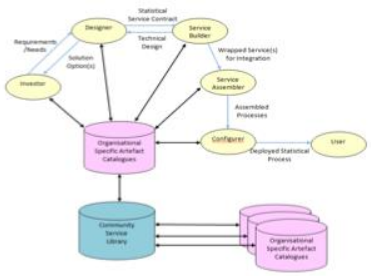
Confidentiality on the Fly - Status

Item	Status
Service Concept	Complete
Service Definition	Complete
Service Specification	Complete
Service Implementation	Complete
Proof of Concept validation (Test)	80% complete
Distribution and Support	Incomplete

Lessons

Lesson #1 - Roles

- The CSPA role model provides a clear and useful suite of differentiated roles
- The role of each Agency's catalogue and the need for an international catalogue is clearly identified
- The Assembler and Configurer roles will be of increasing importance in the future
- Agency feedback from PoC activities demonstrates various levels of maturity in this shift



8 Statistics Canada • Statistique Canada

Lesson #2 - Resourcing

- CSPA Proof of Concept activities were "Idea Phase" R&D activities
- It can be challenging to secure time and effort with the competing focus of "production" priorities
- Necessary expertise could be difficult to secure in predictable quantities
  - Service builder and reusable "software component" experts
  - Production solution assemblers
- Have we "qualified" as "mainstream" ?

9 Statistics Canada • Statistique Canada

- Role clarity has improved, but we continue to miss "real customers"
- Resourcing continues to be a challenge – "Idea Phase" projects, part of Capability development but not part of "Production"
  - Won't be until components are "designed in"

Lessons

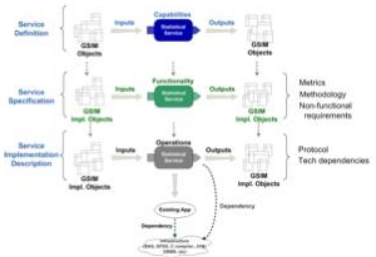
Lesson #3 – Legal Issues

- Assets may not be "given away" – taxpayer-funded assets
- "Right to use", "Right to modify", "Right to distribute" may be freely granted
  - Must currently be administered
- Secondary monetization may be viewed as a source of supplementary income
- New co-developed solutions are treated differently – shared assets
- Plenty of examples in the GPL, OpenSource world that work
- Liability for errors ?
  - E.g. OpenSSL !

10 Statistics Canada • Statistique Canada

Lesson #4 Design & Implementation Models

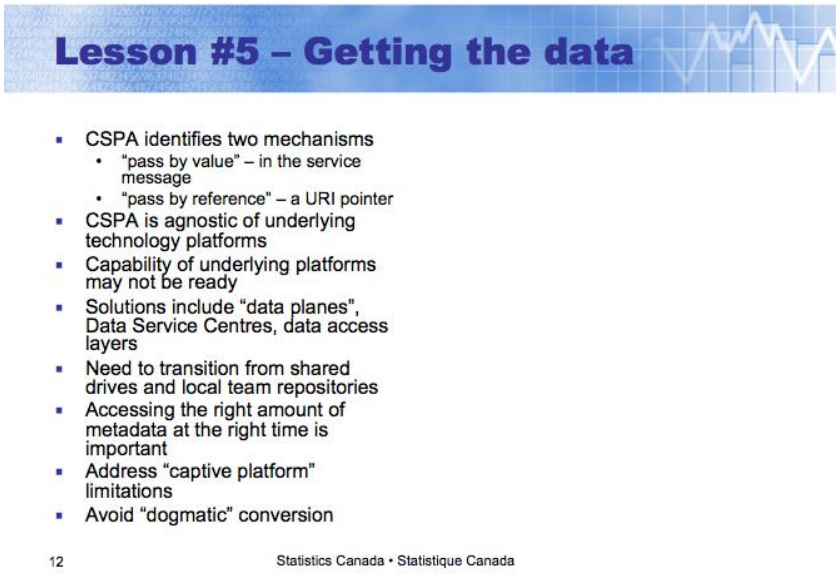
- GSIM is an effective conceptual design model for service definition
- DDI 3.1 provided a useful basis as an implementation model
  - Potential for "overkill"
  - Gaps in representation
  - Need for DDI 4
- "Rule languages" are not addressed but are an opportunity
  - Identified at the Rome sprint

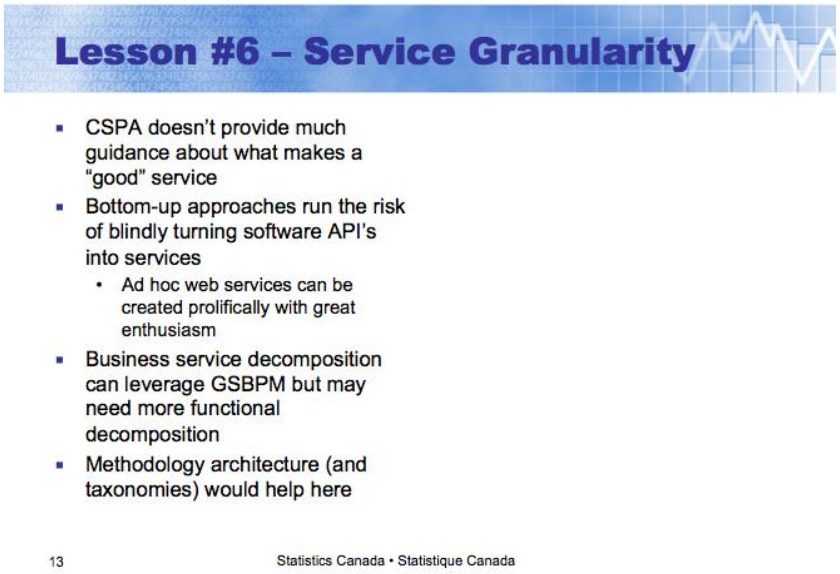


11 Statistics Canada • Statistique Canada

- Legal issues have not been an issue as the component has not been "released" yet
- GSIM an effective design model – implementation model still missing

Lessons



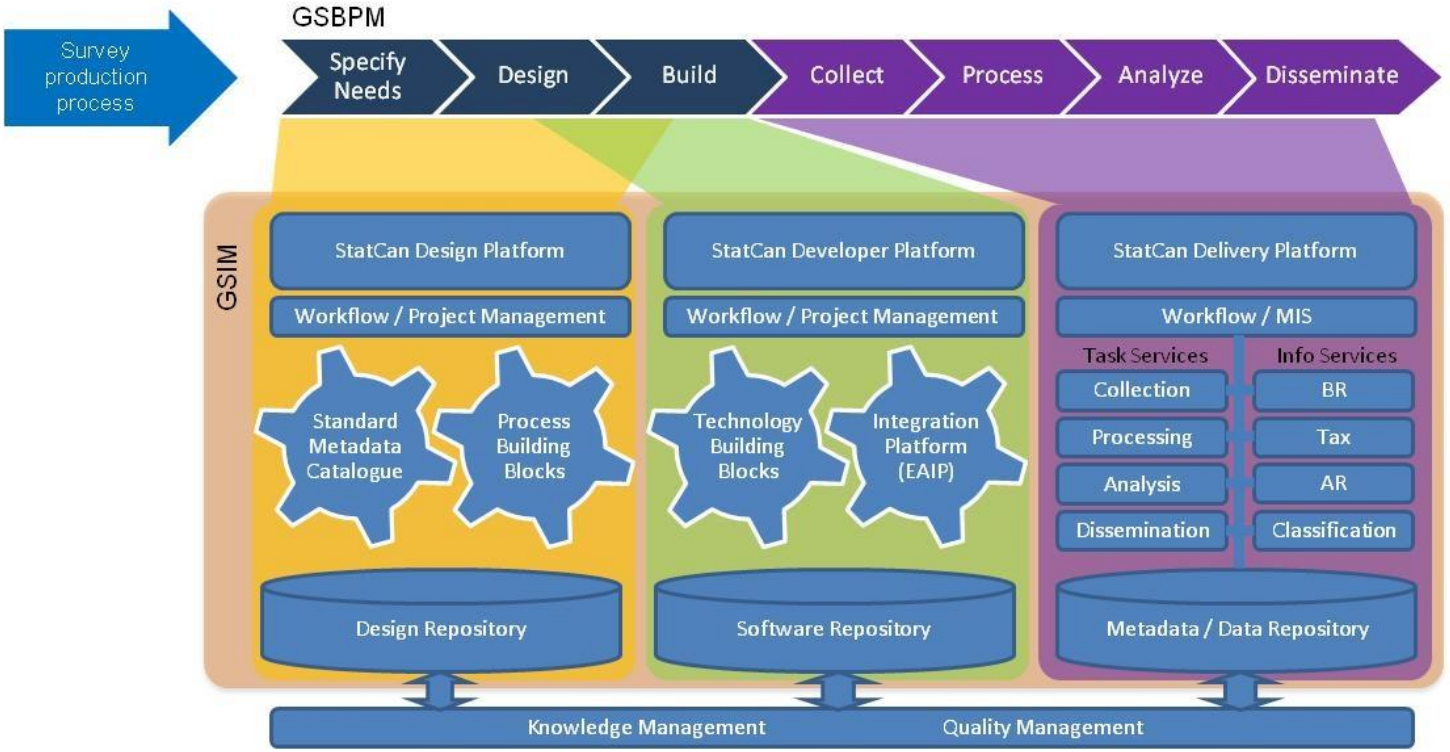


- Data and persistence posed challenges – understanding and designing in solution for local environment – CSPA specifies "pass by reference / value", but this remains a gap
- We had challenges establishing the right level of service and interfaces – require a business capability model and methodology architecture to address this

Lessons – Collaboration

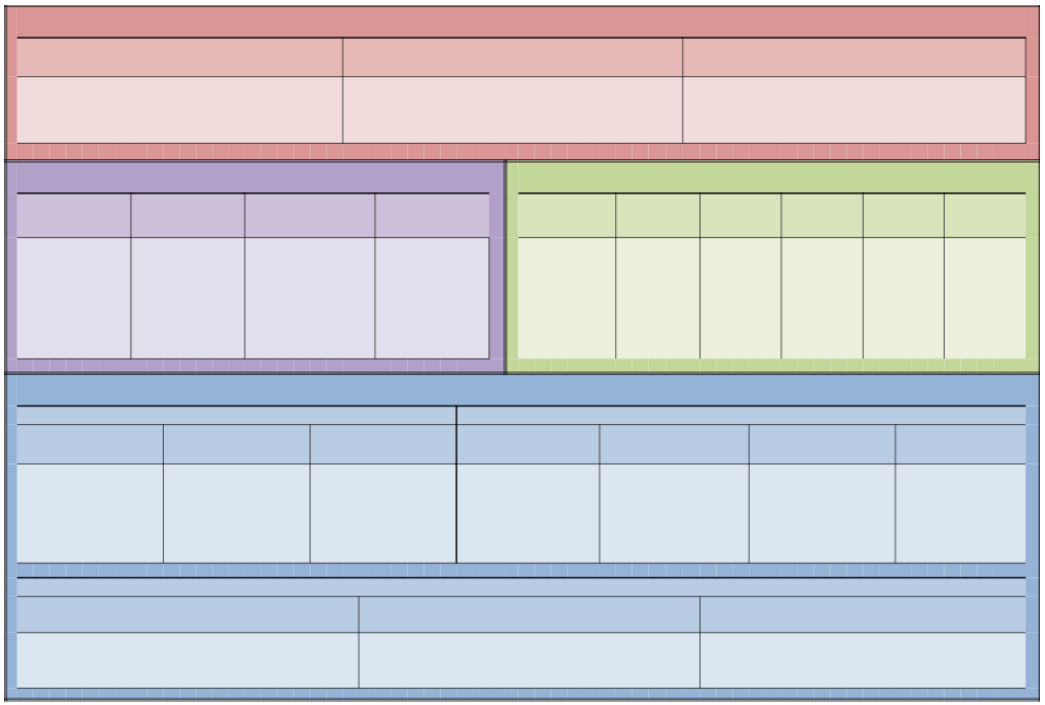
1. Project management is important to ensure momentum and progress
2. Agile model is likely to be most effective
   - Feature backlog drives implementation
   - Multi-disciplinary team required
3. Collaboration environment around Agile is necessary to accelerate progress
   - E.g. Jira-Agile, GIT / Stash
4. Supporting documentation and knowledge sharing requires extra effort across different geographies, corporate cultures, disciplines
5. Silos continue to exist within Statistical Agencies amongst IT, Methodology, Subject Matter
   - Hinders collaboration and more rapid evolution

Concept of Use



- Collaborative component-based (data, metadata, process, method, software) statistical production capabilities deployed within a National Statistical Organization context

Context of Use - GAMSO



- Our Strategy defines our collaborative component-based approach for agility, flexibility, efficiency
- Our Capability program identifies and oversees delivery of new components to the international and internal communities
- Our Corporate Support ensures that necessary communications, collaboration, technology, and legal environments are in place to support the effort
- Our Production program is the target within which we realize the business value of working together

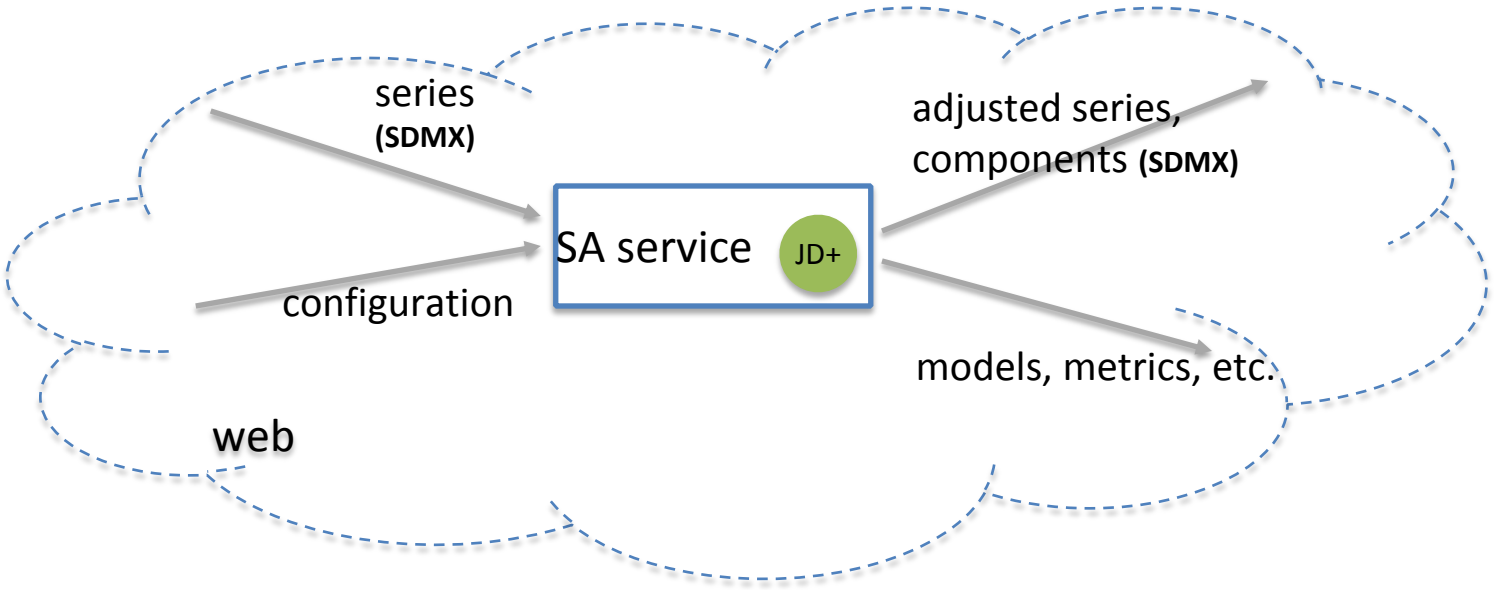
In the Spotlight

Build/Assemble - Seasonal Adjustment (INSEE)

Franck Cotton

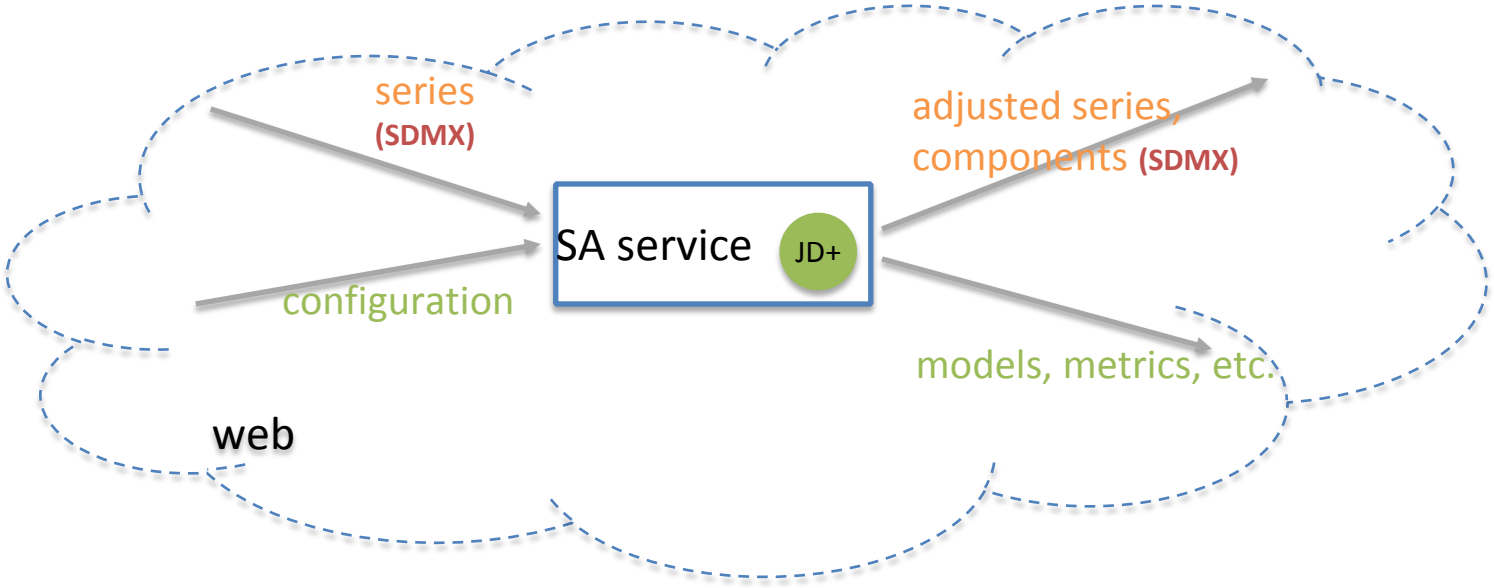
franck.cotton@insee.fr

What was planned



Design	ABS, INSEE, SNZ
Build	INSEE
Assemble	ABS, INSEE

What we did



Design	ABS, INSEE, SNZ
Build	INSEE
Assemble	ABS, INSEE, CBS?

What happened?

- Lack of project management
  - Everybody is busy and has other priorities
  - International collaboration did not work
- « Assemble » is hard
  - More rooted in the organization
  - Needs planning, time, money
- On the whole
  - We could have done better (and we will)
  - But the journey was worth it

What is next?

- Finish what was planned
- Prepare proper implementation
- Consolidate the organization
  - Service owner
  - Maintenance
  - Link with JDemetra+
- Use the service
  - Internally
  - Support it as part of the ESS CoE

In the Spotlight

Assemble/Use - CBS

Marton Vucsan

mhj.vucsan@cbs.nl

Critical Success Factors

- Interdisciplinary team: IT people and methodologists
  - It is not only the code; it is the people and the processes too!
- Agile way of working: Kanban board, Agile methods
  - Many different tasks, good fit for the method.
  - Very little overhead (yes we do make documentation)
- Joining forces with Italy
  - Thank you Monica (and your colleagues)
  - Cooperating on data editing services
  - Using each others services and testing together
  - Designing and using the same technical interface
- Good support from top management
  - No modernisation without it

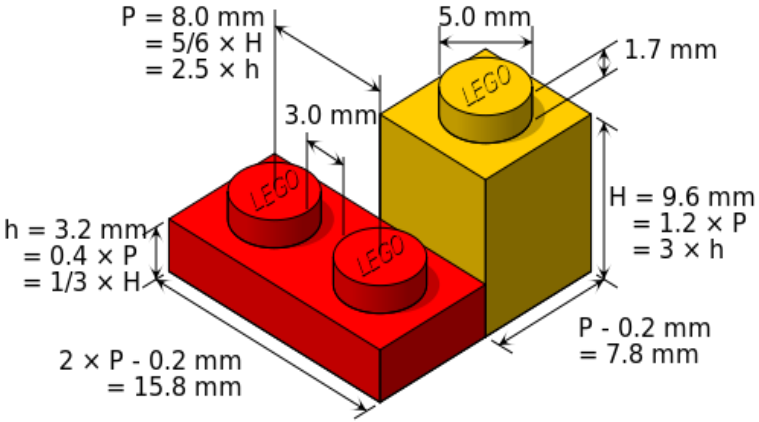
Outputs

- Products:
  - 3 services
  - Technical architecture for implementation
    - REST web protocol
    - JSON table schema for physical communication
  - Design methodology and replacement for BAD/MAD template encouraging and defining CSPA compliant design
  - Local Catalogue
  - Administration procedures embedded in regular application administration
  - Deployment procedures (thou shall not underestimate!)
  - Transition strategy

Findings & Recommendations

Embedding CSPA in your NSI…

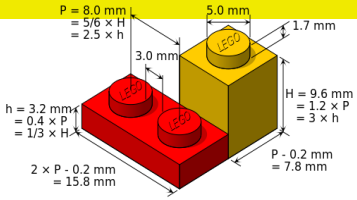
Design



Build

Use

Governance



Design

- It begins with the project portfolio…
  - We need to think in GSBPM terms and CSPA steps and identify project opportunities
  - Generic Model for Statistical Organisations (GAMSO) and Business Capabilities
- Methodology comes next:
  - Design template encouraging designing the methods in clear CSPA compliant process steps
  - Reuse of methods already available as CSPA services; browse the catalogues!
  - At the end CSPA services to be re-used do need to implement shared statistical methods!

Build

- Doing new things is inspiring; costs time
- A shared development environment would help
  - We must find better ways to support the different environments like shared (possibly Cloud) testing and collaboration environment
- Authorisation and security
  - Is a pain, and not very mature
  - Needs to be implemented in separate NSI specific data-services (front: CSPA, back: any)

Build

- We have closed development environments
  - Services are modern -> new tools! (IT dept.: hmmm)
  - Many new tools expect web access during compile
  - Importing executables “needs getting used to”
- New disciplines
  - Creating VMware compatible images (or the like)
  - Deploying services on application servers
  - Orchestration or using “hollow apps”

Use

- Reuse is a way of life (no, no, not like that)
  - Top management should be actively promoting
  - Business case has local costs but central benefits
- Design concessions for reuse
  - It is not our normal mode of operation
  - Needs to become part of our culture

Use

- Portfolio management and investment planning cycle:
  - shared planning and investment should give a priority boost
- 8 implemented services cannot be easily re-used within production if internal environment is not designed for reuse
  - more focus on internal NSI environments i.e. NSIs should work in parallel to the CSPA work to implement SOA architectures

Governance & Administration

- A global and a LOCAL catalogue is needed
  - Promotion from local to global is a process
- Services have a lifecycle
  - Technical and functional administration needed
- Normal application incident processes change
  - Bugs can be local or global
  - Emergency fork procedures should be implemented
- Adjust work and IT processes (i.e. CMDB)

CSPA V1.0 – Governance Proposal

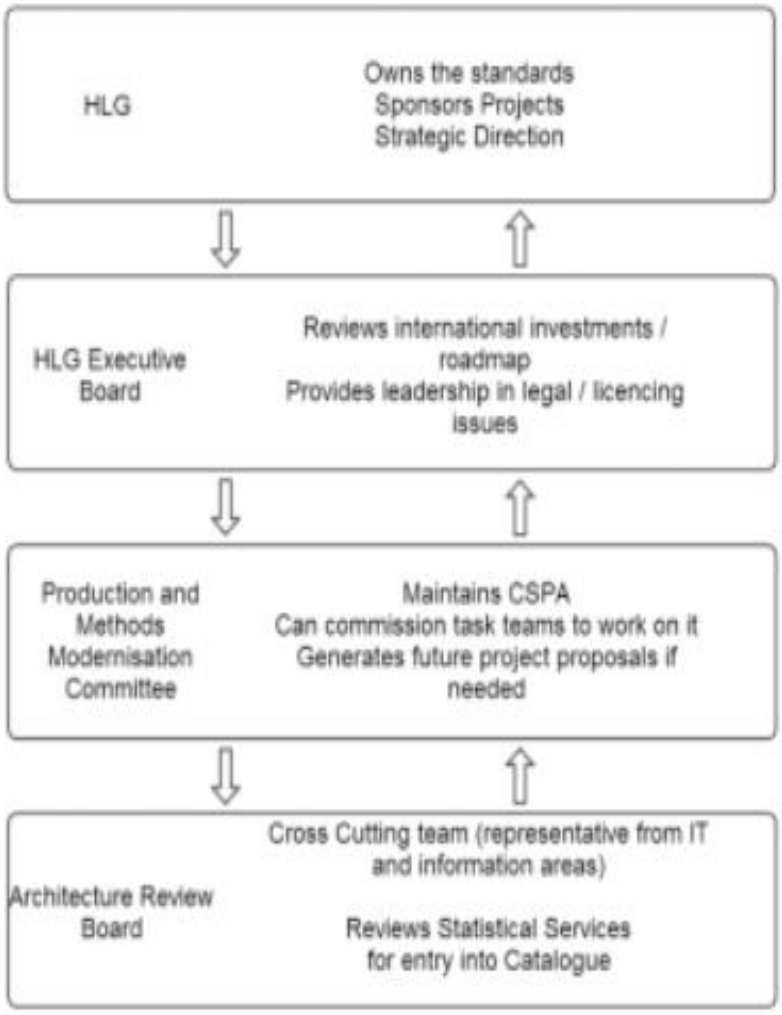


Figure 10: A proposal for CSPA Governance structure

Needs to expand responsibilities to effectively seek collaboration opportunities?

Needs to expand responsibilities to effectively support implementation?

We need our governance to support implementation

Architecture Working Group

The Vision

Awareness Raising & Consensus Building on CSPA

Fostering & Sponsoring Research

Creation and disseminating International Best Practices & Lessons Learned

Capacity and Capability Building

Becomes Reality

CSPA Stewardship

Articulating Global Best Practices & Knowledge

Strengthening Local Implementations

- Maintain CSPA Reference Architecture
- CSPA related Catalogue content approval
- Support Implementation
- Issue management and resolution

- Identify research opportunities to aid implementation
- Joint research with standards bodies
- Technology research to aid implementation

- Publish to Global Artefact Catalogue knowledge base
- Collaborate across Mod Comms for Production and Methods and Standards to ensure alignment

- Training programs and toolkits
- Establish and support technical implementation community
- Develop CSPA implementation champions

CSPA Guidance, Training,

GSIM Implementation Tools

Collaboration Tools

Research and Test Area

Technical Implementation Communities

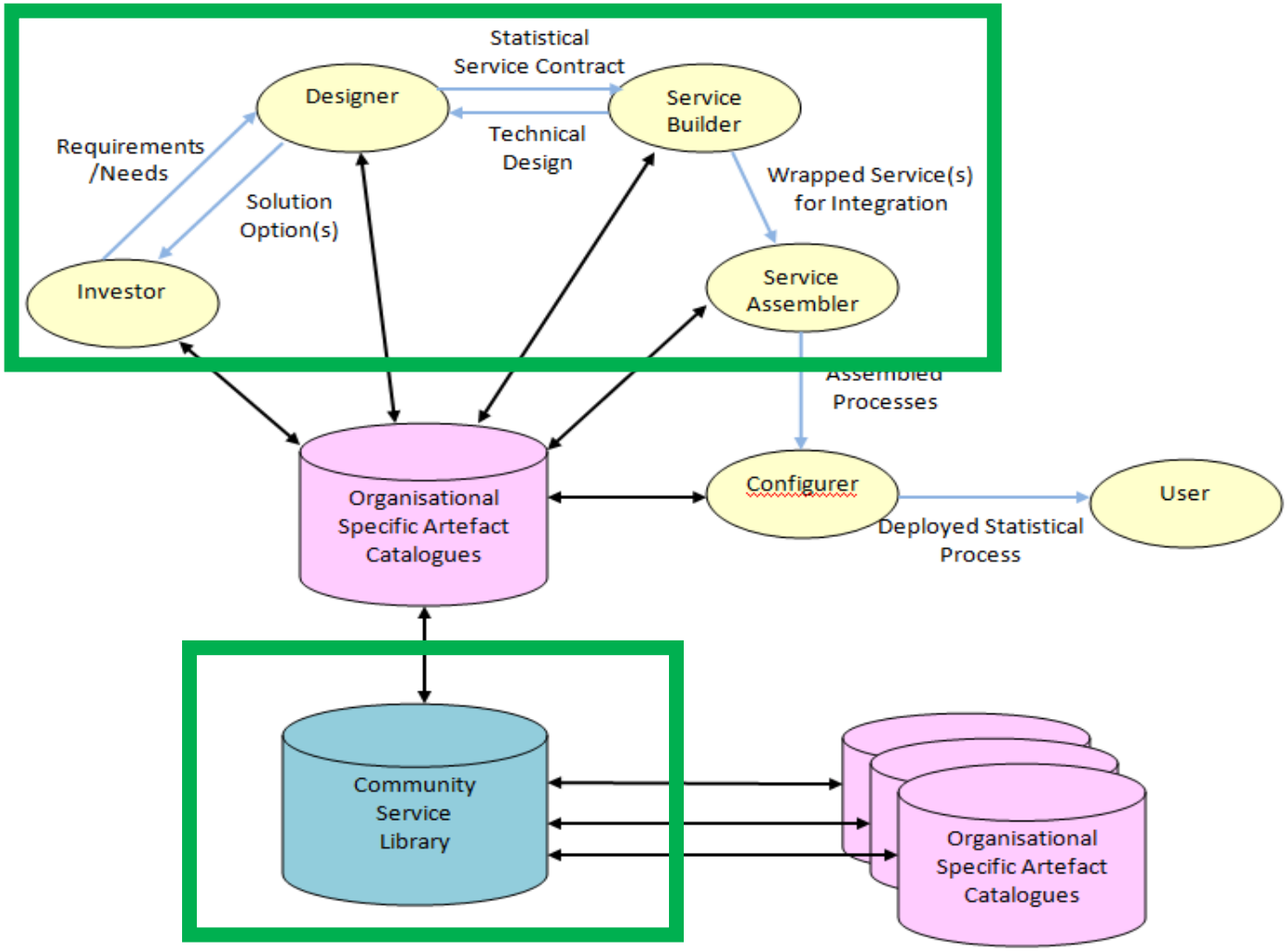
Implementation NSI's, NSO's and Statistical Organisations

Technical Coordination Committee (Proposed)

Next Steps

The Future is Here…

We Need to Shift the Focus



Alignment with ESS Vision 2020

- The alignment of the continued work with CSPA (plus GSIM and GSBPM) and the continued work with implementing the new ESS Vision is a key priority.
- CSPA standard is one of the key standards to structure the statistical services needed to implement the architectural building blocks.
- From a CSPA-perspective, the continued work with implementing the ESS Vision is an excellent opportunity to test the CSPA-standard on a larger scale and to further develop the standard.

Other Opportunities

- Use of CSPA framework in Working group on software sharing as a part of the International Census Forum group.
- Methodology-based alliances like ‘Confidentialise on the fly’ working group leveraging work done in Confidentialised Analysis Service team
- Large redevelopment projects in NSIs…
- External Partnerships (NZ Ministry for Business Innovation and Employment, Norwegian Social Science Data Services)

Practical Issues - Checklist

- Known (from 2013 project)
  - Licensing issues
  - Exchanging software (Import/export) is cumbersome.
  - Installation of software not built for you can be tricky and requires support.
  - CSPA-service could be applied in production only if support is available
- New (2014)
  - Lack of alignment of NSI projects/portfolios to facilitate top down planning
  - Lack of capability for local integration support particularly on the technical level
  - Insufficient governance arrangements to support implementation of CSPA services in full production

Steps Toward Success

- 2013
  - CSPA v1.0
  - Proof of Concept
    - Paper
    - Videos
  - Example services
- 2014
  - CSPA v1.1
  - 8 Implemented Services
  - Supporting Tools
    - Global Service Catalogue
    - Technical Repository
    - Guidelines, Training materials, SW legal framework, agreed process with ModCom Standards for GSIM standards
- 2015?
  - CSPA v1.2
  - Services in full production
  - Production support
    - Technical coordination
    - Shared development environment
  - Support for investment planning and transition to ModCom Production and Methods

Thank you!

Executive Board, Modernisation Committees

2 Working Groups

1 Project Manager

7 Implementation teams

55 individuals