



STATISTICAL INFORMATION SYSTEM COLLABORATION COMMUNITY

2016 WORKSHOP HIGHLIGHTS REPORT





THE AGILE DATA FACTORY: CREATE AND COMBINE DATA EXPERIENCES







We are pleased to present the OECD's Statistical Information System 2016 Collaboration Community (<u>SIS-CC</u>) workshop highlights report.

This year's theme was '**The Agile Data Factory: Create and Combine Data Experiences'**, with a total of 24 presentations over 6 sessions. More than 50 participants representing 20 organisation's attended the workshop over the 1st two days, of whom many expressed very positive feedback. The workshop was also an occasion to welcome the International Labour Organisation ILO as the newest member of the Community.

All presentations are available on the <u>workshop web site</u>, including the workshop agenda. In addition to the presentations you will find a selection of <u>photographs</u> taken during day 1 and 2.

Introduction

Welcome by Omar Baig, OECD 🔄; Keynote by Martine Durand, OECD 🔄; Agenda by Eric Anvar, OECD 🕑

This workshop took place in the context of the <u>5 year 2014-19 strategy</u> and, as such, helped the community to take the next steps to in delivering on the directions set by the Strategic Level Group (SLG):

- More Open Data Dissemination Developing an Open Data Strategy;
- More Innovative Data Dissemination Enabling New Data Experiences;
- More Industrialised Data Dissemination Streamlining Data Processes;
- Implementing International Standards.

At the 2015 workshop we introduced the concept of SIS-CC as part of a broader Open Innovation ecosystem, with the idea of establishing more connections to the broader data ecosystem, projecting ourselves beyond .Stat as a dissemination platform and progressively extending our footprint as a Community, to other areas of co-innovation and co-production of solutions and practices.

This year's topic, SIS-CC as an Agile Data Factory, was a continuation of an idea that came up in last year's workshop. Our emphasis this year was to look more inwardly into the functioning of the community, and how we can reinforce it and further operationalise the concept of open innovation in the context of SIS-CC.

This agenda, as underlined by Martine, is in tune with the SDG agenda, as the release, in full open source, of reusable web components supporting data dissemination, could be a strong contribution by the Community to this agenda, as illustrated also by the LMIS project led by the ILO in the Caribbean region.







SIS-CC as an agile data factory

To visualise the SIS-CC as an agile data factory, four dimensions were considered:



- Information Architecture: Leverage data standards and data structure harmonisation to facilitate data flows, data aggregation, reuse and repurposing. This has been an area of focus for the SIS-CC since its inception, with constant investment in <u>SDMX</u> implementation in partnership with Eurostat, especially on the output side of .Stat and now more on the input side as well (notably, to support data aggregation mechanisms that are becoming mainstream across all SIS-CC organisations).
- 2. Application Architecture: Evolve from monolithic software architecture to a componentised one. This evolution is in continuity with the evolution towards a web service centric architecture, enacted and implemented in 2014-15. With the decision to go fully <u>CSPA</u> compliant in 2015, the SIS-CC is in the process of delivering further in that area, on the output side to start with (in support of new data experiences notably, the forthcoming UIS data portal and the renewed .Stat browser), and on the input side as well.
- 3. **Operational Model:** The SIS-CC application life-cycle management process has been strengthened over time, with development teams already working according to the Agile SCRUM model, but in loose connection; the question for the workshop participants to reflect on was: How can the various teams work in a more connected and synchronised way through a <u>distributed Agile SCRUM model</u>?
- 4. Business Model: The current SIS-CC business model, reviewed and consolidated when the <u>5 year 2014-19</u> <u>strategy</u> was formulated, corresponds to 'closed open source': Software code that is open for community members and free of charge (no licensing fee), but is not accessible to non members (who are not contributing to the Community activities support, coordination, promotion and product development). With the evolution towards a componentised architecture, there is an opportunity for the SIS-CC software to be more easily usable by a larger audience, without disrupting the pooling of resources that the community enables. The question for the participants was to explore precisely that option.

It was recognised, during the workshop, that:

- Given that dimensions 1 and 2 have already been extensively explored in previous years, with directions set and confirmed, the objective achieved during the workshop with regard to these two dimensions was to take stock of progress made and lessons learnt, and refine the plans and roadmap.
- The focus on day 2, especially during the final session, was to explore options for the Community to evolve along dimension 3 (Operational Model to synchronise better the different development teams in member organisations through distributed agile mechanisms such as SCRUM-on-SCRUM) and dimension 4 (Evolution towards a hybrid open source model, leveraging the componentised architecture that is being progressively rolled out).







Session 1: SIS-CCommunity in action

The SIS Collaboration Community workshop is in its 6th year and continues to go from strength to strength. With the collaboration on statistical activities between statistical organisations gaining importance, both at national and international levels, this session accomplished the following: Reformulated the SIS-CC 2015-19 strategic plan and its level of completion; Demonstrated how the use of a common platforms is being jointly developed to enable cost-effective innovation in a minimal time benefitting a number of organisations; Provided an opportunity for each Community member to share their experiences in regard to successes, challenges and future directions and provided an opportunity for partners and organisations interested in joining in to share their views and expectations from the Community.

Presentations:

Business Case for joining the Community Edgardo Greising, International Labour Organisation (ILO) 🕑

Member Project updates

Australian Bureau of Statistics Italian National Institute of Statistics National Bank of Belgium Statistics Estonia UNESCO Institute for Statistics UK Data Service DATA S



2015 achievements and looking ahead to 2016/17 by Jonathan Challener, OECD

Architecture evolution by Chris Beer, ABS, and Jens Dossé, OECD 🕑

Day 2 – The Agile Data Factory: Create and Combine Data Experiences

Keynote: Unity is strength by Laurent Carbonnaux, Ekito 🕑

What a beautiful adage! But why do we have to work together, what is the purpose and how can we work together when people are not together? Agile strongly recommends having co-localised team members. What to do when that's not physically possible, when we're not in the same office, in the same country, or the same organisation? How are others doing it? What can we learn from the Open Source community? Some suggestions: A common vision, emotional colocalisation, creating your own method.









Session 2: The Agile Data Factory – Key Concepts



Agility implies a level of flexibility, being able to adapt quickly and continuously. A factory typically is the assembly line, being a set of predefined rules and blocks, centralised and standardised to enable a continuous delivery mechanism to churn out a well-defined product. When these two are combined, production allows for quick and easy delivery that is able to be reconfigured and evolve to meet the changing demands of the business, while mutualising and capitalising on shared software assets.

This session brought forth the foundation of the Agile Data Factory, highlighting the building blocks that are in place today within SIS-CC, the processes and practices that have been adopted, the technology choices, and how this enables organisations to quickly develop new data products.



Topics and Speakers:

SDMXSource: A common component architecture across SDMX compliant software Nadia Vlahova, Eurostat 🗹

Implementing SOA: The .Stat componentised architecture Jens Dossé, OECD, and Chris Beer, Australian Bureau of Statistics 🗗

The Open Source Software Community paradigm Rowan Wilson, OSSWatch 🕑

The CKAN Open Source Software Community Use Case Sebastian Moleski, OKFN 🕑

Data Intelligence Service Centre (DISC): A data centric approach to IT Gerard Salou, ECB

Session 3: Quick fire panel

Following on from the previous session, a panel consisting of the speakers from Session 2, discussed the key points, and provided the answers to the audience questions. The panel was facilitated by Gerard Salou from the European Central Bank.









Session 4: The Agile Data Factory – Data Processes and Integration

Integrating various data streams, by leveraging standardised production processes and systems, we are able to bring together the statistical data production and data analysis, and in-turn allow users to interact with various data sources through tools that they are familiar with. One example of this is Excel. This session brought forth a number of examples at the input end of the Agile Data Factory, including: Data Lifecycle Management; Back-office providers and Data Aggregators, which feed the input part of the chain to create and combine Data Experiences.



Topics and Speakers:

Building a modular data management ecosystem Gyorgy Gyomai, OECD **I**

SDMX-RI and Registry: A vision on metadata driven data dissemination Francesco Rizzo, Italian National Institute of Statistics

The .Stat Data Wizard: Facilitating access to data sources and easing the data production processes Samuel Pinto Ribeiro, OECD 🗗

Session 5: The Agile Data Factory – Creating user experiences

Developing data experiences on demand by combining different interfaces allows for a high-level of flexibility and a more targeted delivery, thus helping to meet the ever increasing and versatile demands from end-users. The SIS-CC is a very good example of how to build on a strong foundation to create data services from a single source, integrating a flexible production line that continuously delivers through a joint collaborative effort. This session brought forth a number of examples of products that demonstrate concretely how the Agile Data Factory can be used to combine reusable components to create data experiences.

Topics and Speakers:

The .Stat Data Experience Factory Jonathan Challener, OECD 🕑

Reusable Web Components Project Mark Falvo, UNESCO Institute for Statistics **G**

The many interfaces to .Stat: ILO LMIS project Wei chen Lei, International Labour Organisation 🕝

An academic experience at the UK Data Service Susan Noble, UK Data Service 🗗









Session 6: How can SIS-CC become a network of agile data factories?

This session provided an opportunity for participants to break into smaller groups to discuss specific topics and help to build a vision for the wider SIS-CC Agile Data Factory using the building blocks presented in the previous sessions. The session was chaired by Brian Buffett, UNESCO Institute for Statistics. Break out sessions were animated and reported on by leading experts in the fields of Open source community and Agile development cycle animation.

The ideas, questions and insights gathered fed the members' only part of the workshop (on days 3, 4 and 5) that should lead to a proposition of evolution of the SIS-CC model to its stakeholders.

This session focused on two key dimensions:





Informa Architec Interconnection of these teams are today loose

(SDM)

Introduce a **distributed model** through mechanisms (such as **SCRUM-on-SCRUM**)

Dperational Model

SIS-CC work in an **open source model**, from a licensing perspective

Accessible only for Members contributing to the Info

Evolving towards a **hybrid open source model** becomes an achievable and interesting option to consider













FIND US ONLINE @SISCC.OECD.ORG