OECD
BETTER POLICIES FOR BETTER LIVES

SIS-CC
Statistical Information System Collaboration Community

SUSTAINABLE DATA FOR DEVELOPMENT

BUILDING AN INCLUSIVE DATA ECOSYSTEM

STATISTICAL INFORMATION SYSTEM COLLABORATION COMMUNITY WORKSHOP

SIS-CC WORKSHOP 2017

03-04 APRIL 2017
OECD PARIS, FRANCE

Highlights Report

We are pleased to present the OECD's Statistical Information System 2017 Collaboration Community workshop report. All presentations are available on the workshop web site. In addition to the presentations you will find a selection of photographs taken during day 1 and 2.

Key points:

- This years theme was "**Sustainable Data for Development - Building an inclusive data ecosystem"**, exploring new models for a broader alliance to support developing countries in building data dissemination capabilities,
- The discussion, guided by the recent research by Paris21, was articulated around the dimensions that define the pursued model: open source, extended community support, component architecture and extended data lifecycle coverage,
- More than 50 participants representing 20 organisations attended the workshop over the 1st two days,
- The workshop was also an occasion to welcome INS Tunisia as the newest member of the Community.

Introduction

Day 1 of the 7th Annual Community workshop got underway with Omar BAIG, Head of Digital Knowledge and Information Service at the OECD welcoming Member organisations, Partners: Eurostat; UNICEF, organisations interested in the Community activities and considering to join: Pacific Community; INEGI; CSO Ireland; Bank of Norway, technology partners: TCS; Helmes, experts in the area of open source community practices from OSSWatch, and colleagues from Paris21, who joined in this year to help explore options for a new community model.

The **keynote** on day 1 given by Martine Durand, OECD Chief Statistician and sponsor of the SIS-CC, highlighted the importance of the work of the Community to enable the use of a common platform providing consistent implementation of statistical standards such as SDMX to facilitate data exchange, and allowing for the creation of joint statistical outputs. The Community continues to align to a 5-year strategy with overarching governance through the Strategic Level Group providing mechanisms to pool resources and ensure alignment to each organisation's strategic direction and the wider international initiatives.

Martine went on to say how the Community can play an important role in assisting with the implementation of the new Sustainable Development Goals (SDGs), endorsed by all UN States in September 2015. Data and indicators are playing a big role in the implementation and monitoring of progress towards the SDGs. The OECD is supporting the United Nations in ensuring the success of the 2030 Agenda for Sustainable Development by bringing together its existing knowledge, and its unique tools, data and experience.

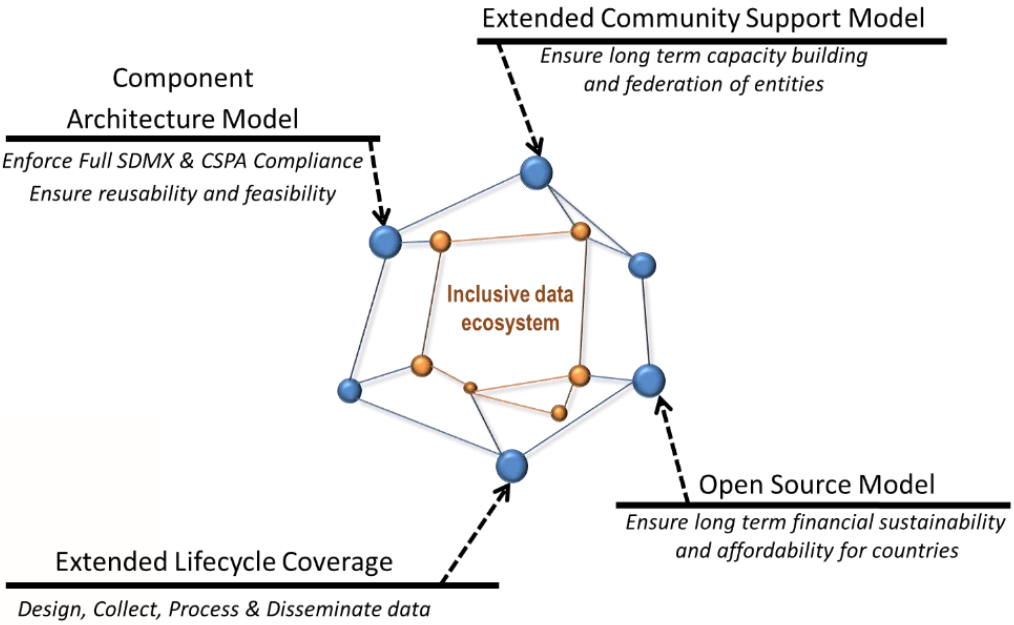
The main theme of the workshop was guided by the recent research by Paris21.

Johannes JÜTTING, Secretariat Manager, Paris21 provided the keynote on day 2 looking beyond dissemination and to Capacity 4.0. The adoption of the Sustainable Development Goals (SDGs) within the context of an evolving data ecosystem is challenging our understanding of statistical capacity. Traditional roles of National Statistical Systems (NSS) stress data production. More and more capacity development efforts should focus also on data sharing and measuring policy effectiveness. User communities are demanding data and are poised to change the role of civil society. Statistical literacy needs to be encouraged; and diversifying how statistics are communicated, the public can be a powerful force for change and development.

Efficient technology and standards have catalysed the process of bringing data users and producers together and facilitate proactive reporting and monitoring of national, regional and international development priorities. While development partners have provided significant support and resources to help enable modern means such as "data portals", a lack of coordinated efforts have provided hindrances to developing autonomous NSOs supported by national budgets. It is not enough to produce and implement new tools and processes but the new environment requires reinvigorating our commitment to bring about real change and sutainability in statistical capacity and our roles in promoting the shifting paradigm. Among them are the need for greater harmonisation of international partners, integration of country statistical processes and the assumption of country fiscal responsibility to search for cost-effective solutions and self-reliance in a sustaining regional setting. Statistical capacity development in this backdrop, demands radical rethinking. It must bring about transformation that is generated, guided and sustained over time from within countries.

Harnessing the Data Revolution well, National Statistics Offices are poised to gain following the Capacity Development 4.0 framework, where overarching driving forces are: digital, integration and partnerships.

Having set the stage for day 2 and the sessions to follow, participants could start to explore the 4 axis addressing many of the questions raised by Johannes's key note and those by the Paris21 research paper.



Day 1

Session 1: SIS-CCommunity in action

The SIS Collaboration Community workshop now in its 7th year, continues to go from strength to strength. With an increasing importance of the collaboration between statistical organisations, both at national and international levels, this session reformulated the SIS-CC 2014-19 strategic plan and the level of completion of it; demonstrated how the use of a common platform enables cost-effective co-innovation benefiting all members; provided an opportunity for each community member to share their experiences in regard to successes, challenges and future directions; provided an opportunity for partners and organisations interested in joining in to share their views and expectations from the community.

Presentations:

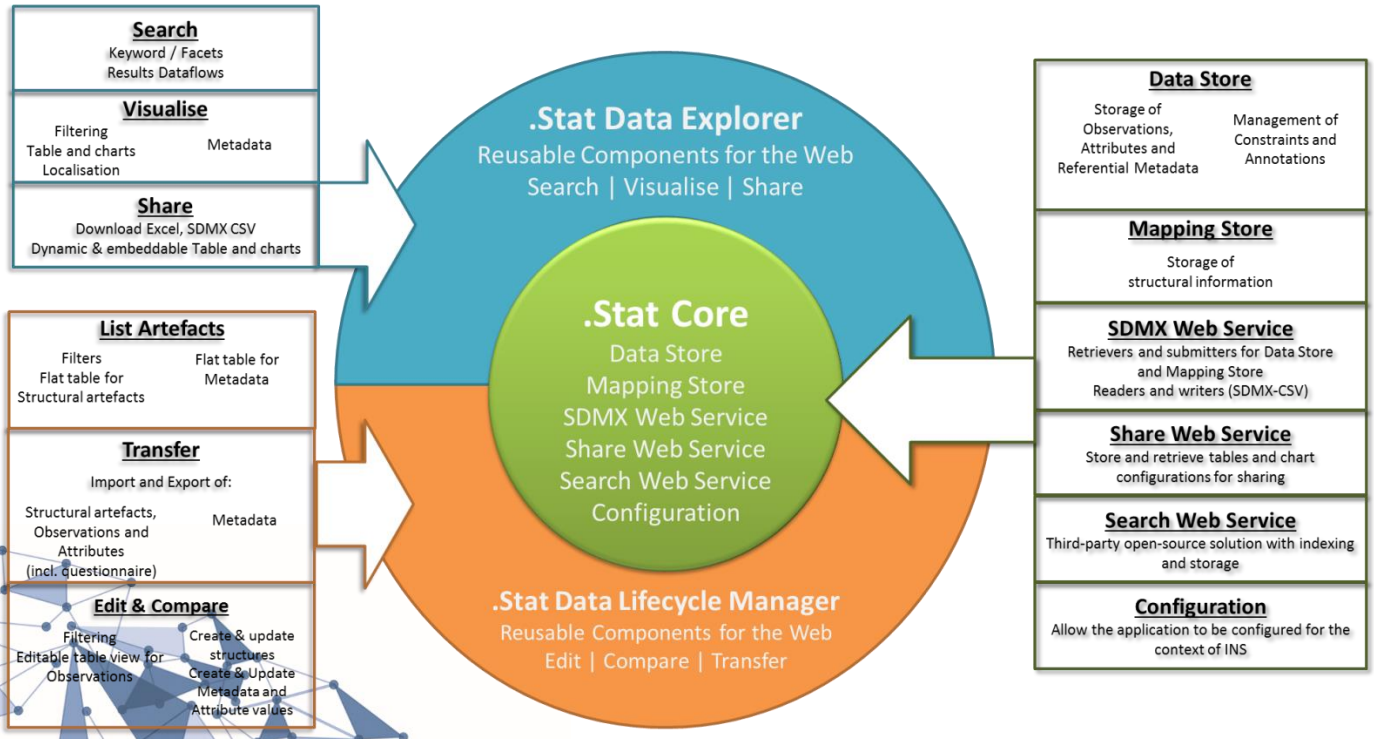
Business Case for joining the Community
Kamel Abdellaoui, Tunisian National Institute of Statistics

Member Project updates
Australian Bureau of Statistics
International Labour Organisation
Italian National Institute of Statistics
National Bank of Belgium
Statistics Estonia
UNESCO Institute for Statistics
UK Data Service

Architecture evolution by Jean-Baptiste Nonin, OECD

2016 achievements and looking ahead to 2017/18 by Jonathan Challener, OECD

New .Stat Suite functional overview



Day 2

Session: Component Architecture Model

Enforce Full SDMX & CSPA Compliance and ensure reusability and feasibility

'As is': Initiatives so far have been designed according to 'monolithic' architecture paradigms. This has led to a low enforcement to standards but more importantly, has limited the possibility for countries to adapt platforms to their specific needs and context (or do so at the cost of 'forking' from the mainstream product). This has also led to one central entity only being able to usefully contribute to the product development, restraining the capacity to contribute to innovation from other entities.

Target: With the proposed model, platforms will be built based on a set of components (building blocks), interfaced with each other according to standard protocols. This requires an important initial investment to set up the proper technical framework – and we are at a point in time when, from a maturity level, this is possible and is being done in the context of the SIS-CC project – and once this is done, countries can assemble the different building blocks much more freely and without too much effort to build tailored data portals, topical or regional data explorers, in a fluid and agile manner. This architecture approach, complemented with the proper technical capacity building, enables to mobilise teams from all around the world in a way that is efficient. This approach is also a way to synergise developments done by other organisations and avoid reinventing the wheel: 25% of the future .Stat suite will be directly sourced from Eurostat open source developments, to which SIS-CC also directly contributes, under Eurostat intellectual property.

Presentations:

SDMX standardisation collaboration & interoperability by Nadia Vlahova, Eurostat

Sustainable .Stat Componentised Architecture by Jens Dossé & Bruno Urban, OECD

Break-out Session Word Cloud

Session: Extended Lifecycle Management

Design, Collect, Process & Disseminate

'As is': Initiatives so far have been focusing on the final dissemination product ('data portal'), with limited or no consideration for the complex back office processes for designing the datasets to be disseminated, aggregating the data from multiple sources, preparing it and enriching it with metadata, according to release management workflows (including embargo and archiving requirements). Also, initiatives so far are not intrinsically SDMX native, which leads to a dual work by data modellers. Another limitation observed is the inability to properly connect the platform with internal data processing platforms, leading to a lot of manual overhead to manage data flows.

Target: With the proposed model, a broader part of the data lifecycle will be covered (and will be further extended in the future, thanks to the componentised approach which will allow to capitalise on progressive contributions by multiple countries and entities). The project should help resolve the limitations observed today, by providing a set of back office components that can be assembled in a way that is tailored for the specific context and processes of each organisation, thanks to the componentised approach. The SDMX full compliance (on the input side, and not only the output side), as well as connectors with main market tools, will greatly facilitate integration with other data processing platforms in use, as well as excel and other similar tools. The project will also help dissemination teams in statistical organisations manage in a much more cost efficient way the different data sources they need to aggregate and set up the proper workflows to ensure quality assurance.

Presentations:

Data Portals: A view on deployment, design and technology by Rajiv Ranjan, Paris21
A sustainable and robust statistical dissemination infrastructure by Kamel Abdellaoui, INS

Break-out Session Word Cloud

Session: Open Source Model

Ensure long term financial sustainability and affordability

'As is': The official statistics dissemination market (to be distinguished from the broader open data govt market) represents a too narrow and specialised market to ensure sustainability of a purely private approach to the market. Several attempts in the past and more recent ones have been sanctioned by bankruptcy and, eventually, huge loss in terms of assets and investments by organisations. The world of official statistics is also one where a very strong culture of international collaboration and standardisation exists since decades, which leans itself very well to open source culture and approaches, as well as strong potential for reusability (the functional requirements being very close, and aligned with standard specifications). These two elements combined (narrowness and specialisation of the market questioning the very possibility of a private player being able to sustain on this market on the long term; culturally favourable environment for open source approach with very similar needs), in the context of the global agenda for SDGs which calls for options that are affordable for lower income countries, makes open source as the right model to address the challenges of financial sustainability.

Target: The model proposed is open source, with a core of active or lead members ensuring financial sustainability over the long term by sharing the cost of product development and maintenance, while making it available to the broader set of users for free. This model has proved to be sustainable for the SIS-CC (.Stat) project over the past decade (although the code was open to members only so far), and will continue to be so going forward with the code becoming open source – as it is recognised that the main reason for contributing financially to the community is the capacity to co-produce and influence the product roadmap. With the advent of a broader community of core members, with funding capacity, the financial sustainability is even more secure, while being combined with affordability for the larger group of countries. With the componentised architecture, useful in-kind contribution to the product development will become more and more accessible, which will further increase the capacity to mobilise resources even if they are not financial

Presentations:

DevInfo community model and future opportunities by Claes Johansson, UNICEF

Roadmap to an Hybrid-Open-Source model by Rowan Wilson, OSSWatch

Break-out Session Word Cloud

Session: Extended Community Model

Ensure long term capacity building and federation of entities

'As is': Having a product with the right features and properly maintained is not sufficient. You need the capacity to promote and inform potential users and facilitate their understanding of the product and how they can use it in their context – this activity is a fixed cost which increases as the scope of the market addressed extends. More than that: you need a capacity to train users, train also contributors in the member organisations. With the advent of SDMX, it's more than just training users on a tool, it's about supporting them in improving their capacity to design new data products, reengineer their data models and processes. Initiatives so far have been driven centrally (one central entity covering one limited part of the previous, with no network lever effect), which has led to strong limitation in meeting the demand for more skills and knowledge sharing in the broader sense (and not limited only to training on

Target: It is proposed to leverage the community model (where already members are all contributing to various degrees to support each other, instead of having only one central entity retaining all the knowledge) to create a federation of support entities (notably, regional entities or entities with a sectorial specialisation, or the two combined). This model is possible only because the previous conditions are met (component architecture, extended features and open source business model). Any entity contributing actively as part of the core of the community, can develop support activities according to its own brand, its own business model and meeting the demands of its own stakeholders. This model creates the possibility of a vast mobilisation of forces which are the most relevant (and where the resources are): close to the final user, where there is know-how and knowledge of the final users' requirements and constraints – as opposed to a central entity deprived of this knowledge and the resources to engage in human intensive capacity building activities.

Presentations:

ILO and CARICOM project by Edgardo Greising, ILO
Building SDMX knowledge in countries by Francesco Rizzo, ISTAT

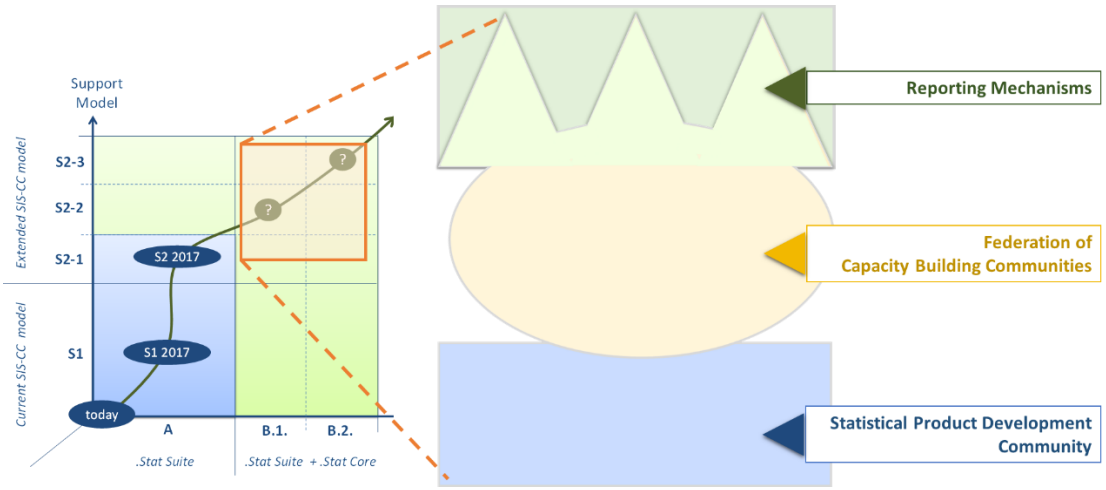
Break-out Session Word Cloud

Session: Final panel

The final session of the day concluded with a panel discussion chaired by Brian Buffett (UNESCO Insitute for Statistics) bringing together a consolidated view from the break-out sessions from each of the session facilitators.

Extended Support proposed model

Following the 2016 workshop, and the endorsement by the Strategic Level Group in the same year, the Community defined a short and medium term path to moving towards an open source model as shown in the figure below 'blue shaded area'.



During day 2 of the 2017 Community Workshop, participants helped to define the vision for a proposed extension of this model to enable the development of a federation of capacity building communities supported by the SIS-CC continued investment and focus on product excellence. This can has been described in the diagram shown below.