SIS-CC 2013 WORKSHOP
'Building products out of our data'

OECD's Statistical Information System Collaboration Community workshop 2013

Highlights report

We are pleased to present the OECD's Statistical Information System 2013 Collaboration Community workshop report.

Thank you to all who participated and contributed to the success of the workshop.

Feedback from those who completed the online survey was extremely positive and it will help in the planning for next years workshop. Thank you for your valuable comments.

All presentations from the workshop, which were made available right after, are still available on the workshop web site. In addition to the presentations you will find links to the Open Data API Initiative submissions as well as a selection of photographs taken during the afternoon of day 1 including the group photo as shown.

SIS-CC Workshop 2013 Participants. Please note: not all participants are shown.

Key points

- We welcomed Statistics Estonia to the community as a new member,
- 53 participants representing 20 organisations attended the workshop,
- Theme was on 'Building products out of our data', a total of 28 presentations over 5 sessions,
- Topics ranged from Community member projects, Tailoring the data experience on .Stat, Business Case for joining the community, Enabling new data products through Open Data, with a special panel session Enabling data exchange through SDMX,
- Two parallel capacity building sessions included .Stat training, and Open Data API initiative for developers,
- Four discussion groups focused on key requirements for the community regarding following features: Embargo, Branding, Localisation, and Graph / Table production,

Overview and outcomes

Day 1: Session (1) Introduction

Presentation topics:
- Welcome and Introduction.
- Community member projects
- Business case for joining the community

The opening and welcome speech was given by Paul Schreyer, Deputy Director, OECD Statistics Directorate, who emphasised how collaboration is a strategic activity for the statistical community and is recognised as such in the aims of the High-Level Group for Modernisation of Statistical Processes (HLG). Paul stressed the importance of collaboration to improve efficiency and move towards the 'industrialisation' of statistical processes. Paul then confimed how this collaboration initiative is of much interest to statistics community. Paul also highlighted how the use of a common platform also enables the implementation of statistical standards such as SDMX to facilitate data exchange, and allows for the creation of joint statistical outputs. This is very much part of the strategic vision for the OECD and of other collaborating organisations.

In 2011 and 2012, the workshop topics were "Laying the foundations for a strong collaboration community" and "Building a collective capacity to enable innovation". Lester Rodriques, Head of Information Technology and Network Services presented the community framework since the first workshop including the vision, and the core elements that support the ongoing activities of the community. Lester highlighted the work to enrich .Stat over the last 2 years as well as how the SIS-CC is influencing much more than .Stat with the initiation of the process to define a new standard format for open data based on SDMX after the 2012 workshop. Lester then gave a brief overview of the agenda for the 2013 workshop under the overall theme of "Building products out of our data". Finally Lester welcomed Statistics Estonia as a new member to the community in 2013.

The session began with a brief project update from the current member organisations; Australian Bureau of Statistics, the European Commission, the International Monetary Fund, Italian National Institute of Statistics, Statistics New Zealand, UNESCO Institute for Statistics, and the University of Manchester on the successes and challenges since 1 year. Presentations demonstrated that collaboration is working well and very good progress being made with a number of public versions of .Stat now live, as well as leveraging .Stat for other initiatives within their organisations.

Eric Anvar, OECD welcoming Alan Vask, Statistics Estonia to the community

The session then continued with Statistics Estonia, INSEE, and the National Bank of Belguim each providing their business case for being interested in joining the SIS-CC and adopting .Stat.

Some of the differentiating attributes of the .Stat solution mentioned were:
- "Rich in functionalities and easy to use",
- "Built for endusers",
- "Very good process automation options", "Modern user interface and functionality", "Good perspective",
- "Develop our own tool will take a long time and a lot of resources",
- "Other similar available tools seem to no longer be improved or have blocking limitations",
- "To be part of a collaboration community".

Day 1: Session (2a) Tailoring the data experience on .Stat
Objective: Define and refine the 2013 workplan

Presentation topics:
- 2012 review and 2013 priorities
- Building a release process with Embargo
- Branding key pain points
- How localisation is managed by others
- Moving from an organisation specific to a localised solution
- Generating charts and tables for multiple outputs

Chaired by Merry Branson, Assistant Statistician, Head of Customer Services Branch, Australian Bureau of Statistics, this session focused on enriching .Stat capabilities in terms of tailoring the user data experience to specific business needs and audiences. Jonathan Challener of the OECD gave a brief review of the global workplan and product evolution in 2012, and then highlighted the key priorities for 2013 as expressed by SIS-CC members.

The session then moved onto more detailed presentations from four key areas, Embargo by Statistics New Zealand, Branding by Australian Bureau of Statistics, Localisation National Bank of Belgium, and the OECD, and Chart and Table generation by the OECD, highlighting the work to introduce an SVG engine and RESTful table generator. Each of these then became the focus of the four discussion groups to help refine the requirements for the 2013 workplan and overall directions for the product roadmap.

Group discussion key points

Embargo implementation options explored
- 2 data sets: 1 public, 1 private with embargoed data. Data set copy switches them when the embargo is lifted. This is triggered manually.
- Load the new data through entry gate when the embargo is lifted. This is fine if it is very fast. It is inadequate when the load takes a long time.
- Load the data into the external database before the embargo is lifted. Restrict access to the data using the embargo view.
- Have > 1 external server with .Stat installed. Pre-load the off-line server, from the internal version of .Stat, before the embargo is lifted. When the embargo is lifted swap the servers. Then load the latest data onto the (new) off-line server before loading the next release of data onto it ready for the next day

Localisation requirements
- Support more than 2 languages
- Support for multiple devices which can have a different interface
- An easy way to translate everything. In the current version of .Stat, labels are stored in different places like database, resource files, XML files. The translation is not a one-off exercise, but should be considered as an on-going process. Labels must be maintained
- The translation of the documentation is a local issue. We should invesitgate if online integrated documentation is possible
- If different languages are allowed, it seems necessary to indicate one language as the default language

Branding
- Centralised Themes and Styles repository
- User guide template although still specific to each organisation
- Move remaining hardcoded references to configuration
- Take advantage of localisation changes to manage labelling

Table and Graph production proposed approach

- Widget approach, based on .Stat JSON web service
- To include metadata (unit, scaling, disclaimer and referential metadata links), automated referencing (source information)
- Query builder and report builder with preview for business users; Guiding business users without strict limitations
- Data selection through restful SDMX syntax, layout parameters (with defaults); Adapt options to type of data
- Style harmonisation throughout domain/organisation through standard CSS
- Ability to 'Get & share' query
- Inbuilt automated analytical features for trend-line, average, variation & growth rates; Role of 0, split of y axis
- Workflow(s) support
- To replace today's tools: Excel chart based on templates (producing jpg or png for web), other statistical tools such as SPSS, STATA, SAS
- Broader solution reusable for paper publication and other contexts: underlying components available through separate installation package

Day 2: Session (3a) Enabling new data products through Open Data

Objective: Present existing and future .Stat Open Data capabilities, and the innovations it enables

Presentation topics:

- Open Data at the World Bank
- UNESCO UIS data portal project
- OECD data publishing with APIs
- Standards and technology to support Open Data services

This session was about reviewing .Stat Open Data capabilities as per 2013 roadmap, and sharing the products that these capabilities enable. Chair Brian Buffett, Head of Dissemination Systems, UNESCO Institute of Statistics introduced the presenters with a brief introduction on how .Stat will support Open Data strategies driven by the SIS-CC members.

Siddhesh Vishwanath from the World Bank opened the session and presented Open Data at the World Bank from the products, managing the various users, and increasing reuse of the World Bank data through a community of developers. Today the APIs have approximately 100,000 distinct users calling them every month. From a fully implemented Open Data environment to a journey that has just began with a presentation by Andrew Barton of UNESCO Institute of Statistics who showed how they will move from a product centric to a task centric environment, with an important integration with the analytical content. A lot of their focus to date has been on visualisations with many examples of different uses and technologies.

Next two presentations by the OECD, both focused on APIs. The first by Terri Mitton covered the publishing side and how the OECD will bring together many sources to create a data portal enabling users to easily find the information they want. The second, by Jens Dossé, looked at the .Stat approach to APIs from a more technical perspective including a brief overview of the developments to date to make OECD public data more accessible to the development community through OData and JSON APIs.

Day 2: Session (5a) Enabling data exchange through SDMX panel session
Objective: Present existing and future .Stat SDMX capabilities and how it can support process automation

Presentation topics:
- SDMX Global Data Structure Definition (DSDs) pilots
- SDMX BOP DSD and SDMX for SDDS Plus NSDPs
- The evolution of the SDMX Reference Infrastructure
- SDMX-RI for data collection and integration in .Stat
- What, why and how of an SDMX registry
- SDMX in .Stat and changes to support the global DSDs

This special panel session was introduced by Trevor Fletcher who enforced the message of the importance of common standards such as SDMX and how it is very much part of the OECD long-term strategy for data collection and dissemination. Trevor also highlighted SDMX in .Stat is vital component for data sharing and open access. In recent times SDMX adoption has increased, and with the globally agreed DSDs, this has brought about new challenges and opportunities for members which this session explored.

Each panel member gave a brief presentation focusing on key aspects of SDMX. The first two pesentations from David Barraclough, OECD, and Gareth McGuinness, IMF focused on the global DSD pilots introducing two new abbreviations SDDS and NSDPs, that is 'The Special Data Dissemination Standard' and 'National Summary Data Pages' which uses SDMX as a standard to publish these under the new SDDS.

Next the focus was on SDMX-RI with a project update by Bengt-Åke Lindblad, Eurostat on the current development path and move to the SDMX 2.1 standard. A .NET version is expected by the end of June with an enhanced version ready sometime in October. Francesco Rizzo, iStat then presented the work to integrate the SDMX-RI into .Stat and how it can be used for supporting data collection in an international context. iStat have a comprehensive architecture vision with SDMX-RI at the center.

Chris Neslon, Metadata Technologies LTD, talked about the role of a registry presenting the Fusion Sandbox as one example of how this can support SDMX needs including a repository for structural metadata, and managing data and metadata sources via 'registrations', which can be used to automate data collection of aid in the discovery of data sources. Jens Dossé, OECD, then presented SDMX in .Stat, highlighting how the first implementation of a fully compatible 2.0 SDMX web service to support the exchange of data has evolved over time, and planned evolutions to support the new challenges that organisations are facing in light of the global DSDs. An important message that was taken forward into day 3 was how far do we go to implement SDMX within .Stat especially considering the large number of artifacts that exist versus those that are already supported.

Special session (1) .Stat training
Objective: Capacity building

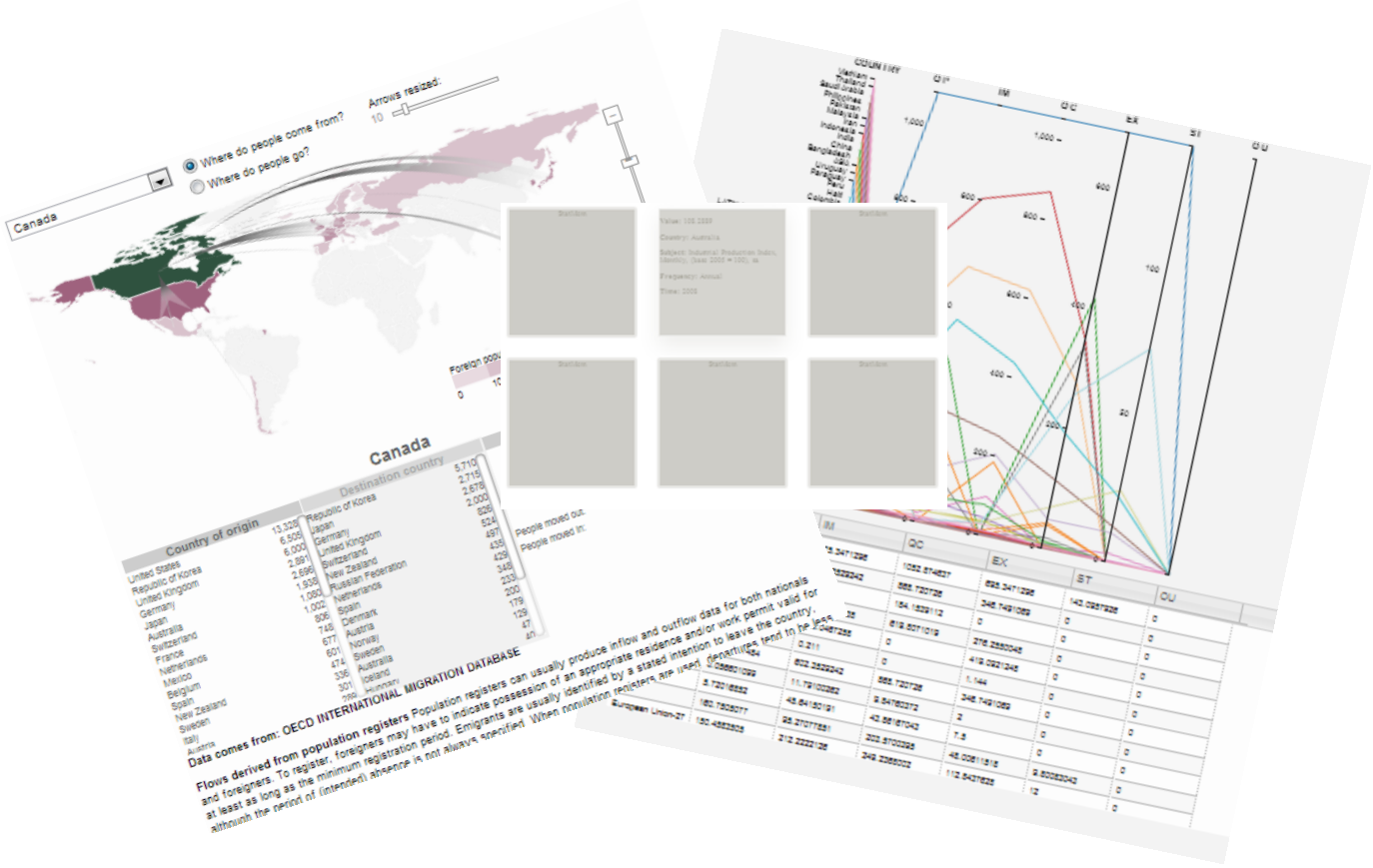
Hands-on parallel session provided interested participants with a detailed overview of the data loading process into .Stat. Conducted by Samuel Pinto Ribeiro from the OECD.Stat support team the session was attended by 10 eager participants either ready to discover .Stat for the first time, or extend their existing knowledge by learning new tricks and tips. Overall the feedback was positive which has encouraged us to hold a similar session in future years.

Special session (2) Open Data API initiative for developers
Objective: Capacity building and open innovation demonstration around data visualisation

This session was the culmination of the Open Data API initiative for developers launched on 27 March 2013 aimed at demonstrating the benefits and the ease of use of the .Stat Open Data web services, and more importantly to gather feedback from developers on this beta version to feed into the final version of the API. The OECD exposed its data through two new beta web services, 'OData' and 'SDMX-JSON', to interested web developers to develop innovative applications based on OECD public data.

Facilitated by Chris Nelson, Metadata Technologies LTD, the workshop was held over 1 ½ days to allow participants to learn more about the APIs, and discuss the different implementations so far including ease of use, available documentation, and wish list for future enhancements. During the morning of day 2, three of the four submissions were presented along with the key points taken from the exercise.

The submissions:
INEGI: http://www3.inegi.org.mx/rnm/sdmx
OECD: https://c9.io/h4mu/sdmx-json-apps/workspace/memory/memory.html
UIS: http://www.uis.unesco.org/das/Country/OdataTest
DRASTIC DATA: http://www.drasticdata.nl/ProjectOECD/



Appendix A – Full list of presentations

Presentation	Presenter	Organisation	Session
Welcome and Introduction	Lester Rodriques	OECD	Introduction
Community member project: Australian Bureau of Statistics	Merry Branson	ABS	Introduction
Community member project: European Commission	Marek Chroscicki	EC	Introduction
Community member project: International Monetary Fund	Gerard Salou	IMF	Introduction
Community member project: Italian National Institute of Statistics	Francesco Rizzo	iStat	Introduction
Community member project: Statistics New Zealand	Jocelyn Morrison	SNZ	Introduction
Community member project: UNESCO Institute for Statistics	Andrew Barton	UNESCO	Introduction
Community member project: University of Manchester	Nick Syrotiuk	UoM	Introduction
Business case for joining the community: Statistics Estonia	Alan Vask	Statistics Estonia	Introduction
Business case for joining the community: INSEE	Danny Kerschieter	National Bank of Belgium	Introduction
Business case for joining the community: National Bank of Belguim	Jean-Pierre Bonnaud	INSEE (France)	Introduction
2012 review and 2013 priorities	Jonathan Challener	OECD	Tailoring the data experience on .Stat
Building a release process with Embargo	Elizabeth Rayner	SNZ	Tailoring the data experience on .Stat
Branding key pain points	Jim Dentrinos	ABS	Tailoring the data experience on .Stat
How localisation is managed by others	Danny Kerschieter	National Bank of Belgium	Tailoring the data experience on .Stat
Moving from an organisation specific to a localised solution	Jonathan Challener	OECD	Tailoring the data experience on .Stat
Generating charts and tables for multiple outputs	Jens Dossé	OECD	Tailoring the data experience on .Stat
Open Data at the World Bank	Siddhesh Vishwanath	Word Bank	Enabling new data products through Open Data
UNESCO UIS data portal project	Andrew Barton	UNESCO	Enabling new data products through Open Data
OECD data publishing with APIs	Terri Mitton	OECD	Enabling new data products through Open Data
Standards and technology to support Open Data services	Jens Dossé	OECD	Enabling new data products through Open Data
SDMX Global Data Structure Definition (DSDs) pilots	David Barraclough	OECD	Enabling data exchange through SDMX panel
SDMX BOP DSD and SDMX for SDDS Plus NSDPs	Gareth McGuinness	IMF	Enabling data exchange through SDMX panel
The evolution of the SDMX Reference Infrastructure	Bengt-Åke Lindblad	EuroStat	Enabling data exchange through SDMX panel
SDMX-RI for data collection and integration in .Stat	Francesco Rizzo	iStat	Enabling data exchange through SDMX panel
What, why and how of an SDMX registry	Chris Nelson	Metadata Technologies	Enabling data exchange through SDMX panel
SDMX in .Stat and changes to support the global DSDs	Jens Dossé	OECD	Enabling data exchange through SDMX panel

Appendix B – Final list of participates

Ms	BRANSON	Merry	Australian Bureau of Statistics
Mr	DENTRINOS	Jim	Australian Bureau of Statistics
Mr	VASK	Alan	Statistics Estonia
Mrs	INNO	Marika	Statistics Estonia
Mr	CHROSCICKI	Marek	European Commission
Mr	SALOU	Gerard	International Monetary Fund
Mr	MCGUINESS	Gareth	International Monetary Fund
Mr	MEISTER	Laurent	International Monetary Fund
Mr	RIZZO	Francesco	Italian National Institute of Statistics
Mr	VIRGILLITO	Antonino	Italian National Institute of Statistics
Mr	SCALZO	Domenico	Italian National Institute of Statistics
Mr	CAMOL	Dario	Italian National Institute of Statistics
Mr	TOWN	David	Statistics New Zealand
Mrs	MORRISON	Jocelyn	Statistics New Zealand
Mrs	RAYNER	Elizabeth	Statistics New Zealand
Mr	BUFFETT	Brian	UNESCO Institute for Statistics
Mr	BARTON	Andrew	UNESCO Institute for Statistics
Ms	PESTINA	Simona	UNESCO Institute for Statistics
Ms	CAMERON	Fiona	University of Manchester (Mimas)
Mr	SYROKTIUK	Nick	University of Manchester (Mimas)
Mr	KERSCHIETER	Danny	National Bank of Belgium
Mr	MOLLET	Marc	National Bank of Belgium
Ms	VAN NUNEN	Petra	National Bank of Belgium
Mr	DE CONINCK	Philippe	National Bank of Belgium
Mr	COMPTON	Terence	Office for National Statistics
Ms	MILIC	Bojan	Statistical Office Slovenia
Mr	LINDBLAD	Bengt-Ake	Eurostat
Mr	CORONADO	Abel	INEGI
Mr	BONNAUD	Jean-Pierre	INSEE
Mr	HENNEQUIN	Benard	INSEE
Mr	BLANC-GARIN	Laurence	INSEE
Mr	GREISING	Edgardo	International Labour Organization (ILO)
Mr	KAUSHIK	Siddhesh Vishwanath	World Bank
Mr	JADUE	Leonardo	Central Bank of Chile
Mr	SECCATORE	Italo	Central Bank of Chile
Mr	NELSON	Chris	Metadata Technology Ltd
Mr	CHAMBENOIT	Eric	Tata consultancy Services
Mr	DUBOST	Romain	InfoCubed
Mr	RODRIQUES	Lester	OECD/ITN
Mr	ANVAR	Eric	OECD/ITN
Mr	DOSSÉ	Jens	OECD/ITN

Mr	CHALLENER	Jonathan	OECD/ITN
Mr	RIVIERE	Bertrand	OECD/ITN
Mr	ZIVKOVIC	Andy	OECD/ITN
Ms	PONTOIZEAU	Katarzyna	OECD/ITN
Mr	KAARTINEN	Petrus	OECD/ITN
Ms	CARTWRIGHT	Susan	OECD/ITN
Mr	PINTO RIBEIRO	Samuel	OECD/ITN
Ms	LEPRON	Karine	OECD/ITN
Ms	MIYACHIYO	Nobuko	OECD/ITN
Mr	SCHREYER	Paul	OECD/STD
Mr	FLETCHER	Trevor	OECD/STD
Mr	BARRACLOUGH	David	OECD/STD
Mr	BROOK	Jonathan	OECD/STD
Mr	OSMAN	Marco	OECD/STD
Ms	TOBON	Carolina	OECD/PAC
Ms	MITTON	Terri	OECD/PAC
Mr	HAMOR	Tamás	TCS/OECD
Mr	VAKULYA	Norbert	TCS/OECD

Appendix C – References

Item	Description
https://community.oecd.org/community/sisccws2013	Workshop website (MyOECD login is required to access)
https://community.oecd.org/community/sisccws2013?view=documents#/?tagSet=2712	Abstracts
https://community.oecd.org/community/sisccws2013?view=documents#/?tagSet=2711	Presentations
http://oe.cd/SISCCWS2013Survey/	Feedback survey (Close 24 May 2013)
http://stats.oecd.org/OpenDataAPI/Index.htm	Open Data API initiative web site
http://www1.unece.org/stat/platform/display/hlgbas/High-Level+Group+for+the+Modernisation+of+Statistical+Production+and+Services	High-Level Group for Modernisation of Statistical Processes (HLG) web site
http://sdmx.org/	Statistical Data and Metadata Exchange (SDMX) web site