

Distr.
GENERAL

CRP.1
13 June 2012

ENGLISH ONLY

UNITED NATIONS ECONOMIC COMMISSION
FOR EUROPE (UNECE)
CONFERENCE OF EUROPEAN STATISTICIANS

EUROPEAN COMMISSION
STATISTICAL OFFICE OF THE
EUROPEAN UNION (EUROSTAT)

ORGANISATION FOR ECONOMIC CO-OPERATION
AND DEVELOPMENT (OECD)
STATISTICS DIRECTORATE

Meeting on the Management of Statistical Information Systems (MSIS 2012)
(Washington, D.C., U.S.A., 21-23 May 2012)

Report of the Meeting on the Management of Statistical Information Systems

Prepared by the UNECE secretariat

1. The Joint UNECE/Eurostat/OECD Meeting on the Management of Statistical Information Systems (MSIS) was held in Washington D.C., United States of America, from 21 to 23 May 2012. Participants from the following countries attended the meeting: Australia, Azerbaijan, Croatia, Estonia, France, Ireland, Italy, Japan, Malta, Mexico, Netherlands, New Zealand, Norway, Poland, Republic of Korea, Russian Federation, Serbia, Slovenia, Sweden, Switzerland, Turkey, United Kingdom and United States of America. The European Commission was represented by Eurostat. Representatives from the following international organizations also attended: Institute for Statistics of the United Nations Educational, Scientific and Cultural Organization (UNESCO), United Nations Industrial Development Organization (UNIDO), International Labour Organization (ILO), Organisation for Economic Co-operation and Development (OECD), International Monetary Fund (IMF), Bank for International Settlements (BIS), and the World Bank.

2. The agenda of the meeting (WP.1) consisted of the following substantive topics:

- (i) Architecture;
- (ii) Streamlining statistical production;
- (iii) Innovation;
- (iv) Collaboration;
- (v) Open data / open government.

3. Mr. Rune Gløersen (Norway) was elected Chairman of the meeting. The preparation of the substantive work was organized by: Mr. Robert McLellan (Canada) and Ms. Branka Cimermanovic (Croatia) for Topic (i); Ms. Jenine Borowik (Australia) and Ms. Christine Wirtz (Eurostat) for Topic (ii); Mr. Matjaz Jug (New Zealand) and Mr. Rune Gløersen (Norway) for Topic (iii); Mr. Marton Vucsan (Netherlands) and Mr. Carlo Vaccari (Italy) for Topic (iv). Mr. Ron Bianchi (United States of America) was Organizer/Chair of the Panel discussion on open data / open government.

4. Ms. Adelheid Bürgi-Schmelz, Director of the Statistics Department at the International Monetary Fund (IMF), opened the meeting and welcomed participants. She stressed the aims of the MSIS group, to provide a forum for statistical information systems managers to share good practices, and to facilitate the implementation of relevant standards across participating countries.

5. Ms. Katherine Wallman, Chief Statistician for the Office of Administration and Budget gave a keynote presentation on the High-level Group for Strategic Directions in Business Architecture in Statistics (HLG-BAS). She stressed that MSIS is very important to enhancing the work in statistical organizations. The CES has played a leadership role in a number of substantive areas throughout the world. This is confirmed by the wide participation in this meeting.

6. The HLG-BAS was created in 2010 to guide strategic developments in how statistics are produced. It comprises the Chief Statisticians of six countries and three international organizations. Modernizing statistical production requires better access to a wide range of data from a variety of sources, to produce higher quality outputs in less time. Collaboration and sharing have to be enabled to move away from expensive systems. This presents one of the biggest challenges to provide quality and fitness for use of statistics. This is a global issue requiring a global perspective. Component based architectures should be based on standardization of production. Statistical organizations could be invited to share their best applications, and pilot projects could be prioritized to fill the gaps. This work might become a natural extension of MSIS.

7. In a special presentation outside the main topics of the agenda, Statistics New Zealand outlined their experiences in the use of Agile approach to software development, in the delivery of the Census of Population 2011/2013. This has been a very successful way of delivering a complex project on schedule and within budget. It has also proved popular with staff. The “Sprint” concept from the Agile approach has also been successfully used to accelerate the development of the Generic Statistical Information Model. However the success of the Agile approach could raise new problems by increasing expectations to unrealistic levels.

RECOMMENDATIONS FOR FUTURE WORK

8. Participants to the meeting identified the following topics for future work:

- Architecture
 - Effects of big data/data mining
 - How do organizations make decisions and how to influence them
 - Define strategy for mobile devices
 - Security – managing cyber threats
- Streamlining statistical production
 - Overcoming legacy system problems
 - Managing organizational change / internal communications
 - Why are we continuing to build separate tools in each organization?
 - Feedback from the CES/HLG-BAS
 - Achievements towards the HLG-BAS vision
- Innovation
 - New technology mini courses
 - Tools for story-telling
- Collaboration
 - Look at the wider data industry – what are our limits?
 - Examples of good collaboration / OECD SIS group
 - Evaluation of tools – why are they used?
 - Roadmaps for collaborative tools
 - End-user requirements
 - Quality / maintenance of tools for sharing
 - Sharing infrastructures
 - Alignment with other expert groups
- Panel discussion

- End-users on what is missing, what they need / external / HLG-BAS members
- Bridging gaps between different disciplines
- Focus groups to tackle key issues proposed during brainstorming session
 - Killer app to demonstrate collaboration.

9. In his summary the Chairman thanked the IMF for hosting the meeting and providing excellent facilities for the work, the MSIS Organizing Committee and the UNECE secretariat for their roles in preparing it. He addressed special thanks to Branka Cimermanovic (Croatia) for her contributions to the Organizing Committee and to the MSIS meetings. On behalf of the participants, the representative of the Netherlands thanked the Chairman for his efficient chairing of the meeting.

10. Subject to confirmation, the OECD kindly offered to host the next MSIS meeting in Paris, France, in the second half of April 2013, where these topics will be addressed.

FURTHER INFORMATION

11. The conclusions reached during the discussion of the substantive items of the agenda are contained in the Annex. All background documents and presentations for the meeting are available on the website of the UNECE Statistical Division.

<http://www.unece.org/stats/documents/2012.05.msis.html>

ADOPTION OF THE REPORT

12. The participants adopted the draft report before the Meeting adjourned.

ANNEX

SUMMARY OF THE MAIN CONCLUSIONS REACHED DURING THE MEETING ON THE MANAGEMENT OF STATISTICAL INFORMATION SYSTEMS (MSIS 2012)

Topic (i): Architecture

Session Organizer/Discussant: Branka Cimermanovic (Croatia)

Papers by: Australia, Japan, Mexico, Sweden, United States, Eurostat and New Zealand

1. The presentation by Australia discussed the Generic Statistical Information Model (GSIM) as a cornerstone of the strategic vision of the High-Level Group for Strategic Developments in Business Architecture in Statistics (HLG-BAS). It facilitates the industrialization of official statistics, with the aim of reducing costs and improving timeliness. It explored what GSIM is (and is not), and how it can contribute to common reference architecture. The process followed when developing GSIM was outlined, including lessons learned regarding approaches to collaborative development which might be adopted by future collaboration activities. An innovative approach was used to progress with the development of GSIM through the use of two Sprint sessions, which were very useful in accelerating progress.
2. The representative of the Ministry of Information and Communication of Japan presented their Optimization Plan of Operations and Systems for Statistical Work, adopted in 2006. This is a strategic plan to promote e-government as a part of administrative reform and aims to overcome the disadvantages of the decentralized statistical system of Japan. They seek to improve public statistical services; improve convenience for survey respondents and security of data; reduce the public statistical workload; reduce cost and to centralize the statistical information systems previously operated and maintained independently by different ministries. The “e-Stat” site provides “One-Stop-Service” that collects and integrates statistical information from the websites of all ministries and provides various statistical services via Internet. It also contributed to providing information about “The Great East Japan Earthquake”. Costs and statistical work load have been reduced and good cooperation among all ministries has been built.
3. Mexico presented the value of having a well-designed systems architecture that is capable of being adapted to several circumstances with little or no modifications, in order to streamline the statistics business process. This architecture allows to describe a complex system at a high level of abstraction. It is a planning and coordination instrument to document knowledge, experiences and best practices. Software product lines consist of a product-line architecture and a set of reusable components. Components are designed in such a way that they can be configured, specialized or easily replaced by providing standardized interfaces. Systems are developed so that they can be reused in different statistical projects and can be shared among organizations. New initiatives and requirements, like open data or development for mobile devices requires systems with robust architectures.
4. The representative of Statistics Sweden described their experience in creating statistical services based on business process models and business information models. He explained how the transition from conceptual models to system logic could be performed, in line with the principles of Service-Oriented-Architecture (SOA) in order to ensure that the statistical services benefit from the advantages of SOA. The Zachman framework provides additional perspectives with the GSBPM and GSIM seeming to fit best at the contextual level. Services would encapsulate fragments of the production process, using GSIM objects to communicate with each other.
5. The National Agricultural Statistics Service of the United States presented their paper on virtualizing and centralizing network infrastructure at a decentralized agency taking into account government mandates and agency business needs. This approach has produced significant cost savings and improved the robustness of the server architecture. It is now possible to access the system from a wide range of devices using virtualized desktops. The changes were cultural as well as technological.

6. Eurostat presented the development and promotion of the plug and play concept as an architectural model in the European Statistical System Joint Strategy. Industrialisation and interoperability are the cornerstones for implementation. This approach builds on the Common Reference Architecture (CORA) and Common Reference Environment (CORE) projects, carried out by groups of European countries over recent years. These projects have provided a common reference architecture for building statistical instruments and production processes based on combinations of a set of components. Implementation of this approach is due to start in 2013

7. The representative of New Zealand presented their Enterprise Architecture Blueprint, the benefits of which include better value for money, increased use of government data, and more flexibility to respond to changing needs. This is a key element of their Statistics 2020 improvement programme. He presented a case study on the integrated data infrastructure. The alignment of statistical systems architectures is based on clear business ownership. This approach allows statisticians to be proactive and directly engaged, and facilitates the management of change.

8. The session organizer identified the three main themes from the presentations, industrialization, sharing and cost reduction.

9. Points raised in the discussion included:

- It is possible to develop components that are reusable across different projects, different platforms and different organizations. However, it is already difficult to do this within an organization, so to do it across several organizations is a major challenge.
- The alternative is to keep building the same components in different organizations, duplicating efforts. It is not sustainable for each organization to try to do everything by itself.
- How to develop an architecture that would be helpful for all. We can share the common components that we need. But we need to share knowledge about common interfaces.
- The potential cost savings provide very powerful arguments, but whilst the costs are similar for all countries, the benefits are greatest for larger countries. Mechanisms to help smaller countries to benefit from the work done by larger countries would be useful
- The CORE and SOA approaches presented by Eurostat and Sweden are fully complementary.
- Governance mechanisms for development projects need to be set up to actively encourage re-use of components, not to just make them available.
- The key is granularity – granular processes – how big or how small. How to strike the balance to make a usable component that is not too big or too small.
- The use of open-source components should be explored further.
- The importance of the leadership provided by the HLG-BAS vision and strategy.
- It is important to bring together experts from different disciplines. HLG-BAS seminars are important in this respect.
- One approach could be to identify which organizations could be centres of excellence in specific areas.

Topic (ii): Streamlining statistical production

Session Organizers/Discussants: Jenine Borowik (Australia) and Christine Wirtz (Eurostat)

Papers by: IMF, Eurostat, United States (2), ILO and Japan

10. The information industry is changing rapidly to one where very short production cycles and instant delivery of information have become the norm. In response, many statistical organizations are examining their production chains with the intention of making them more relevant in this environment. This session focused on efforts of official statistics providers in streamlining statistical production.

11. The representative of the IMF presented a paper on how they are evolving their business model to meet the demands for new data. Challenges in the fields of data collection, validation and dissemination

were highlighted. The Data Gaps Initiative provides data to support analysis of spillovers and interconnectedness. Managing the volume of data with a flat budget has been a challenge, and improvements are needed in respect to data acquisition and validation. New governance structures now support data initiatives and the publication strategy. The key themes have been to streamline, standardize and automate.

12. The presentation by Eurostat described a corporate approach to processing microdata in their organization. The first step focused on internal harmonization and consolidation, through the creation of a generic, modern production chain using state-of-the-art technologies to process microdata, covering the entire workflow from data validation to dissemination. The Generic SAS Tool (GSAST) architecture became operational in 2006-2007. The GSAST platform offers a uniform user interface with integrated metadata handling and clear visualisation of the workflow. It has proven to be useful for several microdata collections. The GSAST Metadata Editor has addressed the problem of the relatively heavy metadata for survey maintenance. However, the possible extension to the data providing organizations presents several challenges, particularly concerning confidentiality and the adaptation of working methods.

13. The U.S. Census Bureau presented the transformations they have made in their household survey operations. Like all successful organizations, there is a need to innovate, become more efficient, flexible, and responsive to customers. Data-driven management demands a new mindset along with new technologies and systems. In view of escalating costs, it became apparent that there was a need to streamline the oversight activities of the management infrastructure. As a result, the Director initiated an ambitious restructuring effort, led by senior managers from both the Headquarters and Field operations. Strong engagement and communication with stakeholders supported the change.

14. The representative of the ILO presented their efforts to streamline data compilation and the dissemination of labour statistics. High maintenance costs, low coverage and problems of comparability between data were some of the most important gaps that determined the urgent need to redesign the system. There is evidence that with their new system there are significant improvements through increased coverage, improved quality, reduced overburden etc.

15. The representative of the Statistics of Income (SOI) Division of the U.S. Internal Revenue Service (IRS) described how they have leveraged technological advances to improve data collection processes. These improvements, which include fully or partially automated data collection procedures, more sophisticated quality verification, and expanded sample and item coverage, have allowed SOI to improve data collection efficiency and enhance data quality, while reducing overall costs. Data are being provided faster, and are being augmented with population data to provide analysts with more comprehensive information. These data have the potential to transform the understanding of behavioural reactions to tax policies and economic shocks. Ultimately these changes have the potential to impact tax, as well as broader government policies.

16. Japan presented their efforts in optimizing business processes and systems of national accounts and support information systems, in line with e-government total optimization policies. This approach supports the regular production of national accounts data, as well as changes such as the introduction of an economic census and the implementation of the 2008 version of the System of National Accounts. Processing has moved from an old mainframe environment to more user-friendly Excel and VBA tools. The choice of Excel was driven by the level of IT literacy and preference of users.

17. Points raised in the discussion included:

- Whilst international organizations described challenges in obtaining data from national statistical organizations, there are clear parallels in the collection of data at the national level. There is a great opportunity to collaborate.
- How to balance the need for user-centric solutions such as the use of Excel, with concerns about system efficiency and security. Japan will develop an asset management system to mitigate these issues.

- The problem is not so much the tools but more about data management practices, system design and change in working methods.
- Strong governance and leadership are essential to change working methods.
- Web-based solutions may be the best approach.
- Corporate rather than local solutions are needed.
- Electronic data collection significantly reduces costs.
- Harvesting data from supplier web sites can help, but is constrained by the lack of standards for publishing data. The notion of setting up a global registry would help, including metadata as well as data.
- Active engagement in the GSIM process will help us to get there.
- If using SDMX you need to develop the tools. Excel could be used as an interface, with macros producing SDMX files.
- Eurostat is working on a common language for data validation, describing validation rules.
- If you want to improve your business you have to control it. You have to have precise definitions and know about your data to be able to make efficiencies in processes.

Topic (iii): Innovation

Session Organizers/Discussants: Matjaz Jug (New Zealand) and Rune Gløersen (Norway)

Papers by: Australia, BIS, United States, Italy and IMF

18. Jonathan Palmer, CIO, IMF, opened the meeting on the second day. He stressed that the environment has never been better for innovation and collaboration. We have excellent communication technology and collaboration tools, which we need to harness for official statistics.

19. The session organizers quoted the definition of innovation from Wikipedia, and identified the most innovative elements from each paper. “Innovation is the creation of better or more effective products, processes, services, technologies, or ideas that are accepted by markets, governments, and society. Innovation differs from invention in that innovation refers to the use of a new idea or method, whereas invention refers more directly to the creation of the idea or method itself” [Wikipedia].

20. The Australian Bureau of Statistics (ABS) presented their activities aimed at accelerating statistical transformation in the context of the information revolution. Challenges include growing private sector competition, changing user expectations and the recruitment and retention of suitable staff. It is imperative to find a way to reduce the cost and time of doing business in order to effectively industrialize and modernize statistical production. Their focus is to move from just being able to survive in the new data industry, to being able to thrive and provide the outputs users want. Re-designed enterprise architecture and re-engineered business processes are based on the greater re-use of tools, modules and components, and the creation of a “plug and play” environment. Their strategy is to accelerate change and achieve savings by concentrating resources on the projects that offer the biggest potential for savings.

21. The Bank for International Settlements presented the data portal they are building with SDMX; a global cloud solution for the statistical community. They have created a sandbox environment in partnership with fourteen central banks. They stressed that SDMX is more than just a data exchange format. Users want good data, a unique place to go to look for this data, and web services to automate procedures. They have proved that many of the building blocks of a full SDMX distributed environment can work in the sandbox. As a result, they intend to continue and expand this approach, as it has proven to be a great tool for learning and understanding “hands-on” the SDMX framework, and it proves that the SDMX vision can be implemented in practice.

22. The US Energy Information Administration presented their work to create highly interactive websites for the dissemination of statistics. Improvements in the current generation of web browsers and an abundance of free or low-cost JavaScript libraries enable the rapid design and development of rich, highly interactive web applications as vehicles for the dissemination of statistical information to meet

user expectations. For example, a rich dissemination platform for electricity data was developed in just ten weeks. Organizations must develop the expertise to create interactive web applications with rich user experiences in order to remain relevant.

23. A representative of the Italian statistical office presented an innovative feature of their recent population and housing census; an integrated web information system, supporting all the activities of the collection process. He drew attention to the strong integration with the census management system, the multilingual support, the metadata-driven approach, and the solutions to improve the quality of data collected. He also described a sophisticated mechanism developed for the assisted coding of textual responses, based on a dictionary, similarity string comparison and automatic ranking. Just over one third of all census responses were provided via this web tool, which will now be used for other data collection activities.

24. The IMF gave a demonstration of their e-library site containing pre-defined data reports based on a unified data model. This approach gives the possibility to create queries across datasets, giving greater flexibility for users.

25. Points raised in the discussion included:

- We should take a look at how commercial industries collaborate and innovate
- How to balance high innovation in targeted areas like the census versus high standardization to achieve savings corporately
- How to persuade big IT vendors to support standards like SDMX: It is difficult to find one that can deliver the kind of complex requirements that are needed
- Is there a need to reconsider Excel as a data collection system?
- Public competitions for users to create applications using official statistics have been successful for several organizations, and can be much cheaper than hiring consultants
- Re-using and sharing requirements could be more effective than sharing solutions. Australia and New Zealand are trying this in the area of data acquisition, looking for ideas on how to expand or replace the Blaise collection tool
- Software is often not designed for sharing. New guidelines on multilingual software development from the Sharing Advisory Board are useful in this respect. Developers of software that can be shared are encouraged to post the details in the software inventory on the MSIS wiki
- We should define more precisely the common requirements, to try to attract the attention of the commercial software industry
- We need to find ways to encourage and speed up innovation, as it is not always a top priority for statistical organizations.
- Providing application programming interfaces (APIs) is important, as this gives others the opportunity to add value to statistical outputs
- How to decide which data to collect, as we can't collect everything
- We should agree on which standards we want to use in the statistics industry, as other industries have already done
- What can we do differently to take advantage of “big data”
- New data provide exciting opportunities to create new products, as well as improve existing ones
- Time to market is very important, we have to speed up our processes through streamlining.

Topic (iv): Collaboration

Session Organizers/Discussants: Marton Vucsan (Netherlands) and Carlo Vaccari (Italy)

Papers by: Eurostat, Australia, OECD, IMF, UNIDO, Italy and Netherlands

26. Eurostat presented their experiences in providing data free of charge on their website, and their experiences of working with Google and with the open data community. Eurostat data are disseminated via the Google Public Data Explorer, and integrated into Google search with Onebox. This has

significantly improved the overall visibility of Eurostat data, resulting in increased traffic to their own web site. Making data available in machine-readable formats using open standards and metadata also enables the media or other data redistributors to easily pick up the data and integrate it into their own specific visualization tools for further dissemination. When data are free, lots of things come for free in return from the open data community. Technology is not generally a problem, but a standardized data model and standard descriptions are needed.

27. The Australian Bureau of Statistics presented ideas about a platform for international collaboration. The aim is to develop common standards and share means of production. We can no longer tailor-make our systems, instead we should have an “assemble to order” approach. Several international collaboration models were outlined, with the example of the smart phone industry providing a promising approach. A platform for innovation and collaboration is provided, based on industry standards, allowing the development of plug and play applications that are connectable and substitutable. In the case of official statistics, this could be enabled by GSIM, and would help to implement the HLG-BAS vision.

28. The OECD provided an update on the activities of their Statistical Information System Collaboration Community, based on shared tools such as StatWorks; MetaStore and OECD.Stat (including OECD *eXplorer*). The aims of this community include co-producing and co-developing state-of-the-art statistical information systems, sharing experiences, knowledge and best practices through multilateral collaboration and building of a collective capacity, and enabling innovation at an optimal cost with all members benefiting from each other in terms of ideas and methods.

29. The IMF presented the inter-agency collaboration efforts in the development of the Principal Global Indicators website. SDMX technical and content standards have been used as enablers to improve timeliness. They have developed a generic infrastructure to pull, transform and load data in SDMX format from the websites of partner organizations. Further enhancements are planned to better exploit the potential of the SDMX standards.

30. UNIDO presented research on collaboration methods and tools for effective multi-national collaboration activities. Nowadays collaboration is unthinkable without suitable collaboration tools, mainly based on Internet technologies. A suitable collaboration tool integrates various technologies into a single application that makes sharing and managing information easier. The presentation focused on web conferencing, identifying the key products and important features. The WebEx service provided by Cisco, and used by the Sharing Advisory Board was presented as an example.

31. Italy presented the final report of the Common Reference Environment (CORE) project. Six national statistical organizations, two universities and several observers participated in the work that began in January 2011 and was completed 13 months later. CORE is not a package but a new way of working. It is a proof-of-concept prototype showing the implementation of industrialized (standardized and automated) statistical processes based on the reuse of IT tools developed on different platforms and by different organizations. The project outputs are fully compatible with the GSBPM and GSIM, and support the plug and play approach.

32. The representative of the Netherlands presented the report on activities and results of the Sharing Advisory Board (SAB). The HLG-BAS Strategy seeks to free up resources for new developments through industrialization, use of GSBPM and GSIM. Products must be simpler to produce; exploit the available data; measure e-commerce and globalisation; use the internet resources; organizational change will be necessary and inevitable. The outputs of the SAB on the MSIS wiki were demonstrated, including the software inventory, principles and guidelines for developing multilingual software, and key indicators for benchmarking IT operations in different organizations.

33. Points raised in the discussion included:

- Many organizations are able to make contributions in terms of staff time for collaboration activities, but financial contributions to secure inputs from external experts is becoming

necessary. Australia has committed substantial amounts of money for the development of GSIM, and the UNECE will investigate the possibility of creating a Trust Fund to support HLG-BAS activities.

- It may be better to invest in integrating existing tools and components rather than developing new ones.
- We need ways to better quantify the benefits of collaboration, and calculate the amount of money saved
- Previous collaboration activities based on both top-down and bottom-up approaches have had limited success. We have to do things differently to get the desired outcome.
- A more strategic approach to collaboration, supported and resourced by top managers will help. Small steps to reach big goals is part of the HLG-BAS strategy
- We may need to think in terms of separate platforms for collection, processing and dissemination, as this could be more manageable than a single platform for the whole production process
- More work is needed on licensing and intellectual property issues
- We need common information models to facilitate the sharing of tools and components
- A pilot project is needed to develop a tool that can provide tangible benefits for all statistical organizations

34. A brainstorming session was held to identify potential future projects for the SAB and HLG-BAS:

- Register of interest in requirements (wants / offers)
- Repository of evaluations
- Quick wins – e.g. dissemination?
- Longer term platform development / architectures
- How the work is done as well as what is done
- How to attract right people / need for right competencies
- More on fundamentals – metadata / paradata driven systems: how to make them work
- Reward / recognition – competition best developer / application
- Schema to coordinate sharing efforts
- Repository of skills and expertise
- Developing on an open source model
- How to reach other countries
- Development platforms
- Stop developing non-shareable components
- Support from top of office
- Clearer cost-benefit analysis – share these
- Inventory of developments / prospective projects – facilitate collaboration not just sharing
- Encourage commercial companies to build on our platforms
- Good practices for software development for sharing
- Open data as a platform? E.g. dissemination?
- Open source start from “best in class” systems
- Collaboration advisory board

- User standpoint – what do they want?
- Sharing infrastructure
- Build 1 global application (editing rules?)
 - Let customers (end users) choose
 - Use global sandbox
 - Agile approach / sprint
- Agile approach – change culture
- Global SDMX registry
- Use blogs etc to spread info
- Add demos to tools to be shared
- Differentiate tools by subject matter
- Crowdsourcing
- Google or others to help
- SDMX support for Google public data explorer
- Collaborate on level of principles not just technical level
- Set right expectations, not diverge from original collaboration models
- Create a product that users want – will give publicity
- Facebook profiles
- Publish success stories
- Try agile approach in other areas
- Get message to wider NSO staff
- How to juggle priorities
- SDMX client visualization app to access any SDMX data source
- Common developers communities on Internet

Topic (v): Panel discussion on Open Data/Open Government

Organizer/Chair: Ron Bianchi (U.S. Economic Research Service, Dept. of Agriculture)

35. The Chair opened the session, explained the four phases and introduced the panellists:
- i) Open Data as an input versus an output, the value chain and legal issues (Marton Vucsan)
 - ii) Relationships among open data, applications users; government interpretation of open data, developed versus developing countries (Carlo Vaccari)
 - iii) Government perspective demands of citizens for open data (Jenine Borowik)
 - iv) Consumer perspective (including introduction of the Smart Disclosure initiative) (Joel Gurin)
36. Marton Vucsan raised issues that he found worrying. How far can statistical organizations go down the road of providing all data as open data? They are part of a value chain which is loose – there are no binding contracts. Suitable arrangements and contingency plans are needed. Legal issues include liability for financial losses by users of statistical open data.

37. The discussion raised the following points:

- User requirements – data or statistics, or both?
- We have to think beyond just dissemination, official statistics are for the empowerment of the people. We don't have a monopoly on interpreting our data
- When input data is dependent on other suppliers you are exposed. Keen awareness is important.
- It is no longer sustainable to rely only on directly collected data.
- Legal disclaimers may be necessary to limit liability to users
- Traditional value chains are becoming more like value networks
- Can we do the work better than the private institutions who are doing the same thing?
- Paid data supply arrangements governed by contracts, can reduce uncertainty in supply chains. Memorandums of Understanding may be more appropriate for public sector suppliers
- Government institutions can lose exposure if private sectors get the data out faster
- Users need to have the metadata or contacts to understand data correctly
- We should focus on what we do best – the work of statisticians and not worry about the rest.
- There is a clear need for increased statistical literacy amongst users
- When staff resources are restricted, giving users the tools to analyse data themselves can help

38. Carlo Vaccari discussed the problems for data suppliers. Whilst open data allows a diversity and richness of interpretation, it also raises the problem of contradictory results, and can make it difficult to distinguish between official statistics and other sources of data. Statistical literacy remains an important issue. Sometimes potential users don't know how to use large open data sets. On the input side, open data may replace administrative data in a few years, so we must prepare our organizations. Open data should also be an important output if legal issues concerning microdata are resolved. We must learn to use new tools for data linking, and new techniques for extracting data from the Web. An international standard for licensing data would be useful.

39. Points raised during the discussion included:

- The UK actively engaged with universities to develop an API for population census data
- Standards focus on data – but also engagement – for example the “five stars” of open data engagement
- We should stop treating our data as small children that need protecting, and let them go.
- How to finance communication and collaboration around open data?
- Statistical organizations should supply some tools in addition to the data
- Too much time is spent on gathering data and not on analysing them
- A collaborative visualization tool may help

40. Jenine Borowik discussed the government perspective on demands of citizens for open data. In Australia, the government formed a task force to look at the issues of open data and open government. There was a clash of cultures in the task force – how should the public service engage with the public? There is now government policy on the role of public servants, who have to be clear whether they are speaking as individuals or as representatives of the government. The Fundamental Principles of Official Statistics provide useful high-level guidance to statistical organizations, and the International Statistical Literacy Project aims to educate users.

41. Points raised during the discussion included:

- Making statistics free of charge is a powerful policy.
- How to measure what is a public good? - What is the value of providing open data?
- A study by John Broughton found that the return on investment ratio from open data was 4 to 1
- Side-effects of. Value to having the tools to access a larger cross-section of information. What is the implication for us and our future survival.
- We are not making products for profit, but open data allows others to make money from our data
- There is a similar situation with open-source software. Open doesn't necessarily mean free
- A “killer app” is needed to provide really easy access to open data.

42. Joel Gurin discussed the consumer perspective. Smart Disclosure is the release of data sets in such a way that they can be useful to consumers, for example ratings tables for services such as health and education. Looking at data by use, rather than by source can facilitate consumer choice. There is a growing interest from third parties to create apps using open data. The Smart Disclosure initiative will soon launch consumer.data.gov as a platform for access to open data sets and tools.
