2018

European Conference
on Quality in Official Statistics

Abstracts

Kraków, Poland, 26–29 June 2018

Welcome to Q2018 Conference

Dear colleagues,

Statistics Poland and Eurostat are pleased to welcome you to the European Conference on Quality in Official Statistics Q2018, which will be held in Kraków on 27–29 June 2018. The event will be preceded by a series of training courses on 26 June.

The Q2018 conference is the ninth consecutive quality conference involving the international statistical community. It will cover a wide range of topics relating to methodology and quality in official statistics.

The conference aims to present various concepts and challenges related to topics such as the implementation of the European statistics Code of Practice and cooperation in the field of quality. It also encourages networking between experts from statistical offices, researchers, academics and international organisations.

This traditional biennial conference brings together experts from all over the world, and has become a key event in the calendar of the European Statistical System.

We expect that our collaboration, an example of best international principles and practices, will significantly contribute to the improvement in the quality of our statistical products.

Dominik Rozkrut
President of Statistics Poland

Mariana Kotzeva
Director-General of Eurostat

Statistics Poland

Statistics Poland is the central institution that implements the tasks of official statistics. The mission of Statistics Poland is to provide plausible, reliable, independent and high quality statistical information about the state of, and ongoing changes to, society, the economy and the natural environment, which meets the needs of national and international users.

One of the tasks of Statistics Poland is to cooperate with international and regional organisations specialising in statistics as well as statistical agencies and offices in other countries, and to conduct research and development activities in the field of statistical survey methodology.

eurostat

Eurostat is the statistical office of the European Union. It is situated in Luxembourg. Eurostat's mission is to provide high quality statistics for Europe.

The development, production and dissemination of European statistics are founded on a sound legal basis, comprehensive methodologies, European and international standards as well as on quality control mechanisms.

Eurostat makes its statistics available to everyone, from European citizens and policymakers to representatives of the media and business people alike.

Scientific Committee

Ms Mariana Kotzeva – Director-General of Eurostat, Co-chair
Mr Dominik Rozkrut – President of Statistics Poland, Co-chair
Mr Emanuele Baldacci – Director, European Commission
Mr Eduardo Barredo Capelot – Director, Eurostat
Mr Czesław Domański – University of Lodz, Poland
Mr Henryk Domański – President of the Statistical Council, Poland
Mr Eugeniusz Gatnar – University of Economics in Katowice, Poland
Mr Enrico Giovannini – Chairman of the European Statistical Governance Advisory Board
Mr Gregorio Izquierdo Llanes – President of the National Statistics Institute, Spain
Ms Sibylle von Oppeln-Bronikowski – International Statistical Institute
Mr Józef Pociecha – Cracow University of Economics, Poland
Mr Aurel Schubert – Director General, Statistics Department, European Central Bank
Mr Andrzej Sokołowski – Cracow University of Economics, Poland
Ms Ineke Stoop – Chairman of the European Statistical Advisory Committee

Programme Committee

Ms Claudia Junker – Eurostat, Co-chair
Ms Monika Bieniek – Statistics Poland, Co-chair
Mr Thomas Burg – Statistics Austria
Mr Agustín Cañada – National Statistics Institute, Spain
Mr Tasos Christofides – European Statistical Advisory Committee
Mr Kari Djerf – Statistics Finland
Ms Mária Dologová – Statistical Office of the Slovak Republic
Mr Krzysztof Jajuga – Wroclaw University of Economics, Poland
Ms Zsuzsanna Kovacs – Eurostat
Mr Jean-Marc Museux – Eurostat
Mr Włodzimierz Okrasa – Statistics Poland
Mr Jean-Pierre Poncelet – Eurostat
Ms Marina Signore – Italian National Institute of Statistics
Ms Maria João Zilháo – Statistics Portugal

Content

Parallel Sessions

Special Sessions

Speed Talk Sessions

Poster Sessions

Notes

SESSION 01 – QUALITY MANAGEMENT AND MODEL IMPLEMENTATION I

Title of abstract: Quality management system in the Agency for statistics of Bosnia and Herzegovina – implementation of CAF model, experiences and future challenges

Fahir Kanlic
Agency for Statistics of Bosnia and Herzegovina

Quality plays crucial role for production of official statistics by National Statistical Institutes (NSIs). Statistically speaking, production of high-quality data requires, in the first stage, use of systematic approach in setting up proper quality management system in NSIs. Modernisation of the processes, technologies and complexity of society are the major challenges for NSIs, in general, especially for NSIs of developing countries like Bosnia and Herzegovina (BIH). Based on adopted Strategy for Development of Statistics in BiH 2020, as first step in the systematic approach to quality management, Agency for statistics of Bosnia and Herzegovina (BHAS) implemented Common Assessment Framework (CAF) model of excellence in 2016 as a tool for Total Quality Management (TQM). The aim of this paper is to present experiences in CAF implementation in BHAS, key activities defined, their implementation as well as effects of CAF implementation in BHAS on statistical system in BiH. This paper will also be focused on impact of CAF findings in implementation of the European Statistics Code of Practice (EU CoP) and other tools (user satisfaction surveys, staff satisfaction surveys, etc.) for quality management system in BHAS. As conclusion, the challenges and perspectives for future work on quality improvement in BHAS and entity statistical institutions, as parts of statistical system in BiH, will be presented.

Keywords: Model of excellence, Quality, Official statistics, Quality management system

Title of abstract: Almost there!

Douwe Kuurstra
Statistics Netherlands

In 2017 we were certified for ISO 27001. In the first half of 2018 we will be privacy proof (in 2017 all statistical processes are already privacy proof). In the second half of 2018 we will be certified for ISO 9001. (I'm confident we will succeed.) In the paper I'll will tell you more about the approach which made Statistics Netherlands succeed and the role the quality department played and did not play.

Background

In 2015 Statistics Netherlands made a roadmap for external certification. ISO 27001, privacy-proof and ISO 9001 were the goals to be achieved in 2017 (ISO 27001) and 2018. The main reason for certification was an external proof to show our commitment with quality of process and product, privacy and security (ISO 27001). The European Statistical Governance Advisory Board emphasizes in their annual report of 2017: 'Trust among users and the public in general is of the highest importance'. Statistics Netherlands fully agrees with this and also with the peers who concluded in 2015: Statistic Netherlands should . . . consider also, in the context of external communication, the adoption of recognized industry standard quality assurance/certification schemes. Trust in the way we work can be achieved when experts from outside the office and from outside

the statistical system give the public and yourselves assurance. By the way, by certification for ISO 27001 and privacy we are also ready for the ESS IT Security framework and the General Data protection regulation.

Keywords: Quality management, ISO 9001, certification, trust

Title of abstract: On the Way to Integrated Quality and Risk Management in a Statistical Office

Jolanta Minkevica, Ieva Pauļuka
Central Statistical Bureau of Latvia

Quality plays a crucial role in the European Statistical System's modernisation strategy "ESS Vision 2020". Quality is a broad concept and covers all activities, as well as the overall functioning of a statistical authority. A Quality management system (QMS) integrates the various internal processes within the organisation and intends to provide a process approach for project execution. A process-based QMS enables the organisation to identify, measure, control and improve the various core business processes that will ultimately lead to improved business performance. As the Generic Statistical Business Process Model (GSBPM) describes and defines the set of business processes needed to produce official statistics, it is logical that the QMS of a statistical organisation follows the GSBPM. At the same time, an organisation needs to make sure that it manages risks so that threats are minimised, and potential is maximised. The Central Statistical Bureau of Latvia (CSB) has developed a GSBPM-based QMS, which is aligned with the organisation's purpose and strategic direction and complies with the standard "Quality management systems – Requirements (ISO 9001:2015)" of the International Organization for Standardization. The QMS also serves as a knowledge base. Statistical production processes are standardised throughout the organisation and are also mapped. Risk management is an integral part of the QMS. The presentation will give insight into the way quality and risk management approaches have been integrated within the statistical institution.

Keywords: quality management, risk management, GSBPM, ISO 9001

Title of abstract: Statistical Process description using GSBPM as a reference – Challenges in a process changing environment

Magda Ribeiro, Maria João Zilhão
Statistics Portugal

Statistics Portugal (SP) has a long tradition of statistical process description, starting with a generic model in 1997. Several editions were revised since then and the latest one was recently updated using the last edition of the Generic Statistics Business Process Model (GSBPM V.5) as a reference. Decisions had to be made to have a handbook in a generic format for all staff in statistical production environment, as well as an adaptation between the GSBPM model and all horizontal processes and internal standards in use by SP. The paper will describe I) the new version of the Statistical Process Procedures Handbook model produced by SP, II) its challenges (strong and weak points) in conception while adapting the GSBPM with SP reality of statistical process implementation and III) the subsequent levels of process description added to GSBPM (namely, geospatial statistical dimension), and much aligned with ISO 9001 Norms with respect to documented procedures.

Keywords: Statistical Process, Process description, Documented procedures, Quality Standards, Quality Documentation

Title of abstract: Documentation System STATDOK – a cornerstone for efficient management of statistical surveys

Tina Šijanec, Petra Blažič, Tina Osvald, Barbara Kutin Slatnar
Statistical Office of the Republic of Slovenia

Systematic and exhaustive documentation of the statistical surveys is one of the challenges that national statistical offices have been facing in the recent years. The widely accepted Generic Statistical Business Process Model (GSBPM) provides a very useful tool to accomplish this task in a standardized manner. The Statistical Office of the Republic of Slovenia accepted the adjusted GSBPM model as the general process model in 2012, which was updated in 2017. On the basis of this model, detailed quality guidelines for proper implementation of each of the phases were prepared. This was the basis for preparing the standardized framework of the internal Documentation System of Statistical Surveys, named STATDOK. It is a response to the needs that arise from the desire for more effective management, transparency of the statistical survey implementation, comparison of surveys and the needs of employees for continuous learning. The paper describes the conceptual framework of the internal Documentation System of Statistical Surveys - STATDOK, implementation of the system and its functionalities. It also presents how the Documentation System will be upgraded in the future.

Keywords: Documentation System of Statistical Survey's Implementation, Process standardization, GSBPM

SESSION 02 – QUALITY MANAGEMENT AND MODEL IMPLEMENTATION II

Title of abstract: Beyond Code of Practice – New quality challenges in official statistics

Hans Viggo Sæbø, Anders Holmberg
Statistics Norway

Globalisation, technology, competition and changes in the political agenda have led to a demand for new statistics, but also provided new possibilities in terms of new data sources and partnerships. However, these developments have challenged statistical institutes' compliance with principles and values described in quality frameworks for official statistics. Such frameworks constitute the backbone of quality management in statistical institutes. The European Statistics Code of Practice is an example developed particularly for the production of European statistics, but there are also similar frameworks for official statistics developed by UN, OECD and regional statistical cooperation bodies. They are all inspired by and build on the UN Fundamental Principles, and have developed gradually or revised to address new challenges. However, having a quality framework is not enough, compliance is, in practice, difficult since there often are trade-offs between such principles. Official statistics must be trustworthy, but at the same time relevant and timely, and this may challenge the independence of a statistical institution. Utilisation of new data sources implies challenges for accuracy and reliability, and meeting competition with partnership may harm equal treatment and confidentiality principles. The paper considers these quality challenges, how they are handled in the existing quality frameworks and could be met by the statistical institutes. Statistical professionalism is a key word in this context. Reflecting on new quality challenges may guide the way forward, on creating and maintaining a culture for continuous improvement in European and national statistics. The starting point of the paper is international, but examples are mainly from Statistics Norway. However, these are believed to be representative for several statistical institutes.

Keywords: Quality frameworks, official statistics, relevance, independence, professionalism

Title of abstract: Experiences in Developing Statistical Quality Frameworks

Michael Colledge
Michael Colledge Statistical Consulting, Australia

Over the last 30 years or so a statistical quality framework has become accepted as an essential part of the infrastructure of a statistical office. It provides a systematic mechanism for ongoing identification and resolution of quality problems, and for maximizing the interactions between office staff. It is a basis for creating and maintaining a quality culture within the office and is a valuable source of reference material for training. It makes transparent the processes by which quality is assured and reinforces the image of the office as a credible provider of good quality statistics. It facilitates exchange of ideas on quality management with other national and international producers of statistics. With this in mind, the UN Statistical Division published a template and guidelines for a national quality assurance framework, which were endorsed by the UN Statistical Commission, and many statistical offices have installed a quality framework. So, what have been

the consequences? The paper outlines the approaches and lessons learned in developing statistical quality policies, frameworks, and guidelines in six developing and developed countries and in three international organisations. It describes the common features of, and differences in, the various approaches. It summarises what the impacts have been, what seems to have worked, and what has not. The paper concludes with a discussion of the relationships between quality management, methodology development, metadata management and risk management.

Keywords: quality policy, quality guidelines, quality evaluation, quality and metadata management, quality and risk management

Title of abstract: Quality and Risk Management at Brazilian Institute of Geography and Statistics: a short report of a work in progress

Ana Cristina Martins Bruno
IBGE Brazilian Institute of Geography and Statistics

The paper aims to present and discuss the risk management model integrated to the quality management model under construction at the Brazilian Institute of Geography and Statistics (IBGE). Quality management and risk management are strategic issues, part of the modernization program linked to the strategic institutional plan 2017-2027. In Brazil, risk management has been strongly recommended by the Federal Court of Accounts (TCU), which has developed a model for assessing organizational maturity in risk management based on best international practices such as COSO, ISO 31000 and British Orange Book. The model proposed by the TCU will be used for risk-based external audits and will evaluate management in four dimensions: capacity, translated in terms of leadership, policies and strategies; preparation of people in risk management; application of these capabilities to processes and partnerships; and evaluation of results and performance of the mission. The model will be adopted by IBGE as a self evaluation reference to identify realistics targets for improvements and produce action plans for developing or enhancing risk capabilities. The work has been carried out by the Governance, Conformity and Risks Committee together with the Quality Committee, that is responsible for the implementation of quality policy and process management, based on generic models GSBPM (Generic Statistical Business Process Model) and GAMSO (Generic Activity Model for Statistical Organizations). Although the project is in the initial phase and has not yet generated effective demonstrable results, it is intended to present the model as a Position Paper, portraying a short report of work in progress under the brazilian Institute.

Keywords: Risk Management; Quality management; Risks Management Maturity Model

Title of abstract: Strategies and approaches for managing risks in the official statistics production: ISTAT experience in the modernization programme

Silvia Losco, Rosario Magro
Italian National Institute of Statistics (ISTAT)

Today institutional organizations are facing pressing and emerging challenges going to meet the speed of technology, the demands for change, necessary to ensure efficiency and competitiveness. Change processes are carriers of risks and opportunities because they are able to create value. Processes must be on time and flexible but it is necessary increase attention to the risk analysis, both at strategic and operational level,

to assure the achievement of the goals. Italian national Institute of Statistics (ISTAT), with the modernization program (2016), has adopted a complete program of change with the aim of evolving the statistical production system from traditional survey models based on the direct acquisition of data from citizens and companies towards a model that uses statistical registers. It is an ambitious program that aims to overcome the „silos" vertical processes of traditional statistical production with a high level of risk. The model provides the creation of an integrated system of registers, of a single logical infrastructure of data deriving from administrative sources, from new innovative sources (Big Data) and powered by continuous data flows. The paper describes the framework adopted by ISTAT to organize the activities of the program, focusing on strategies developed for risk management. At the enterprise level one of the actions implemented to minimize the risks of statistical activities is the identification of seven strategic innovation programs and the adoption of a Portfolio and Project Management (PPM) approach. According to Business Architecture model (BA), at the corporate level risk management is implemented through an organization of the activities in thematic portfolios connected to statistical registers and to service for statistical production. At operational level, the statistical activities organization following a management by project approach that select initiatives and organize the work in phases, with a specific control of risks associated to the single phase.

Keywords: Risk Management, Portfolio and Project Management, Modernization, Strategic and Operational Risks

Title of abstract: Recent developments in the Generic Statistical Business Process Model: Revisions and Quality Indicators

Alice Born - Statistics Canada
Nilgün Dorsan, Deniz Özkan - TurkStat
Therese Lalor - UNECE (Switzerland)
Marina Signore - Italian National Institute of Statistics (ISTAT)

The Generic Statistical Business Process Model (GSBPM) is widely used by the statistical community for a range of different purposes from process documentation and monitoring to training staff. It is currently maintained and updated by the UNECE Supporting Standards Modernisation Group. GSBPM is revised every five years in order to keep it relevant and continue to serve as a common framework for the modernisation of official statistics. The current version (v5.0) was released in 2013. The next revision of the GSBPM started in 2017 where statistical organisations were asked to provide feedback on the model. This feedback has been posted to the public discussion forum on the UNECE GSBPM website, and a group of experts are reviewing and proposing solutions to the issues. The most common issues include how to interpret the model, practical use and application of GSBPM in real life. In 2016, Quality Indicators for all sub-processes of the Generic Statistical Business Process Model (GSBPM) (v1.0) were developed in order to monitor the quality of statistical production. The first version focused on quality indicators for surveys, and complemented the quality management process of the GSBPM. There was a need to incorporate indicators pertaining to administrative data. Therefore, this work was expanded to include quality indicators for administrative data. Version 2.0 of the Quality Indicators for the Generic Statistical Business Process Model (GSBPM) was released in November 2017. In this version, the quality indicators are integrated in each GSBPM sub-process for both surveys and administrative data sources. Another output of the quality indicator work was proposing changes to the GSBPM. This paper describes the proposed changes to the GSBPM resulting from input from statistical organisations and Quality Indicators.

Keywords: GSBPM, quality indicators, administrative data, survey, GSBPM revision

SESSION 03 – QUALITY MANAGEMENT AND MODEL IMPLEMENTATION III

Title of abstract: The Swiss Federal Statistical Office quality, process and risk management system

Milos Schaer
Swiss Federal Statistical Office

The proposed abstract by Switzerland focuses on the new quality, process and risk management system of the Swiss Federal Statistical Office (SFSO) introduced in 2016 as well as lessons learned and next steps in this context. Quality, process and risk management are crucial activities for any organization, especially for statistical offices that have to face many challenges in terms of resources and budget, user needs and technological development. Aware of this reality, the SFSO undertook a major revision of its quality management system between 2014 and 2016. During this work, it was also decided to integrate process and risk management in the same system because they are considered as complementary approaches. The main goal of the system is to support and stimulate continuous improvement of SFSO activities and products and ensure compliance with national or international requirements. This system is now operational and lived through a simple and pragmatic mechanism allowing a better coordination of activities in these three areas as well as a better communication of these topics within the hierarchical line. International standards and tools such as the European Statistics Code of Practice (CoP), the Eurostat Quality Assurance Framework (QAF) or the Generic Statistical Business Process Model (GSBPM) have played a central role in the creation of this new system. This system is itself subject to a constant need for continuous improvement. The use of the Generic Activities Model for Statistical Offices (GAMSO) model and other good practices are all development measures that the SFSO intends to take into account for the future.

Title of abstract: GSBPM and ISO as Quality management system tools: Azerbaijan experience

Yusif Yusifov, Vugar Mammadalizade
State Statistical Committee of the Republic of Azerbaijan

Quality statistical production depends on standardization and quality management of the current statistical system processes. Recent years the application of ISO 9001 international standard along with GSBPM for improvement of activity among statistical institutions incurs certain conflicts and discordance. Several statistical institutions put forward application of ISO and some of them the GSBPM, and cannot identify the more consistent one. State Statistical Committee of the Republic of Azerbaijan follows in its activity the application of GSBPM for systematic management of statistical processes towards delivery quality statistical product to users, and ISO 9001 standard for identification and observance of requirements of quality management system towards obtaining of the quality product. Due to application of GSBPM the creating flexible tool for identification and explanation of processes required for production of official statistics, standardization of management of current statistical processes, and the process mapping was achieved, whereas through the application of ISO 9001 standard the observance of current guidelines and manuals, the application

of PDCA cycle in improvement of outputs was gained, as well as appropriate actions were taken to prevent potential inconsistencies by risk-based thinking. Furthermore significant achievements were gained in the field of implementation of monitoring and improvement measures through conduction of internal audits for completion of GSBPM assessment stage. In our opinion, GSBPM and ISO 9001 standard should not be coincided, on the contrary, they have to be applied in parallel and thus improving quality of official statistics and increasing users' trust should be gained. It is more logical to implement quality and metadata management on the stages and sub-processes of statistical processes through the application of GSBPM. Due to application of ISO 9001 standard the management by the administration becomes easier, furthermore it enables to monitor at what extent the current requirements are followed based on "process approach".

Keywords: GSBPM, ISO 9001, quality management, international standard

Title of abstract: Quality management of methodology for official statistics

Kees Zeelenberg, Winfried Ypma, Peter Struijs
Statistics Netherlands

Methodology is a cornerstone of official statistics and one of the major factors that contribute to their quality. This is reflected in the Code of Practice for European Statistics, which mentions that „Sound methodology underpins quality statistics". However, methodology is hard to explain to the average user, and so we must find other ways to ensure trust in statistical methodology and to convince users of the quality of official statistical methodology, such as independence of the methodological unit, transparency of methods, peer reviews, methodological reviews, and internal quality management within the methodological unit. In this paper we elaborate these various ways. We show how quality management of a methodological unit may be based on the Code of Practice. We show how the various elements work together, and how the whole of these elements may lead to certification of the methodological unit, for example by EFQM or ISO. In 2017 the department for methodology and process development at Statistics Netherlands has been certified according to ISO-9001. As an example, we discuss the various steps that have been taken to achieve this certification. In particular we focus on the quality procedures for internal and external reports, recommendations and briefs; the quality assurance of statistical development projects in which methodologists and business analysts participate; the quality assurance of methodological course taught to statisticians, the internal management of the department.

Keywords: quality management, official statistics, methodology, process development, certification, ISO, EFQM

Title of abstract: Innovation and quality culture in INSTAT

Delina Ibrahimaj, Helda Curma
Institute of Statistics Albania

INSTAT in the recent years have worked on different activities to fulfil the requirements of the European Code of Practice. Quality is part of these requirements, so National Statistics Office must define their quality policy and make available to the public. INSTAT declares that the following principles: impartiality, quality of processes and products, user orientation, employee orientation, effectiveness of statistical processes, reducing the workload for respondents are taken into account when performing its tasks. To maintain the

public confidence in official statistics INSTAT have selected Total Quality Management (TQM) as the general model for quality management, quality assessment and quality improvement. The main objective of this model is to establish a general framework which will ensure that statistical production processes meets the highest standards as regards quality and efficiency. This paper/presentation will provide more detailed information on the quality documentation and provide additional information around the quality reporting for INSTAT products. We will also present how quality culture has progressed since the launch of the first Staff Satisfaction Survey.

Keywords: Quality management, Documentation, Efficiency, TQM

Title of abstract: Integrating process documentation and quality management

Silvan Zammit
National Statistics Office, Malta

With the aim of increasing efficiency and productivity, and, above all, improving the quality of its statistical output, the National Statistics Office (NSO) of Malta embarked on a three-year project for the development and expansion of a quality management framework that covers technical issues and ensures business continuity. The main goal of this project is to integrate data and metadata standards and use them as a template for process documentation. This will in turn provide a framework for process quality assessment and improvement. The first phase involved a stock taking exercise to understand better existing processes across all domains, covering technical and operational aspects. The Generic Statistical Business Process Model (GSBPM) was identified as the best model to describe and define the entire cycle of statistical business processes. Additionally, this model provided a basis to agree on standard terminology and to enable further discussions on developing statistical metadata systems and processes, as well a quality and risk management framework. Such institutional tools were virtually non-existent at NSO. Following a series of consultations, training activities and workshops, a number of key domains were identified and duly documented in line with the GSBPM. The proceedings, together with a meticulous consultation and literature review process, allowed the formation of the Quality Management Framework (QMF). Primarily, we define our QMF as a set of technical guidelines focusing on the design, collection, processing and dissemination of statistical processes. Our main objective is to strengthen compliance with the Statistical Code of Practice on a number of quality-related principles. The process leading to the setting up of the QMF also allowed the creation of a series of specific metadata products for our users. We shall ensure a regular review of both process documentation and QMF, to guarantee the application of modern methodologies and harmonisation across domains.

Keywords: Quality management; Harmonisation, GSBPM, Metadata

SESSION 04 – QUALITY MEASUREMENT AND REPORTING I

Title of abstract: Quality handbook and quality principles – progress and lessons learnt

Mirko Herzner
Federal Statistical Office (Destatis)

The development of Quality Principles for each part of the statistical production process according to the GSBPM is a result of an internal analysis concerning Destatis' fulfilment of the ESS Quality Assurance Framework and was further strengthened by the latest Peer Review. A strategy for the implementation and monitoring of the Quality Principles has recently been developed and tested. After the approval of the Quality Handbook (comprising the above mentioned Quality Principles) in March 2017 by the highest decision-making level of the Statistical Offices of the Federation and the Länder, six pilots have been conducted to assess the compliance with the Quality Principles in different subject matter statistics. A number of improvements have been made to the checklist used for the assessment (like the introduction of filter criteria to limit the total number of items to be answered for the individual statistics or the assignment of Code of Practice principles that are influenced by each of the 342 Quality Principles of our checklist). The lessons learnt cover a wide range of issues including the optimal setup for the assessment workshops (at least two people with good knowledge of the GSBPM) and issues with individual quality principles (not all of them are of equal importance; what does "documentation" involve?). The basic facts concerning the development and design of the Quality Handbook have already been presented at the Quality Conference in 2016 (Q2016). This paper is focussing on the progress made and the lessons learnt.

Keywords: Process quality, GSBPM, quality assurance

Title of abstract: Quality system in a digitalised and modernised statistical system

Grete Olsen
Statistics Norway

Statistics Norway has started a program for digitalisation and modernisation of the production process. By 2020 a new system for some statistics will be implemented, and for 2022 the whole production will take part in the new environment. Will the modernised production system give us more and better metadata on quality from the production? Do we need new indicators for quality and new systems for quality management? How can we digitalise the production of quality indicators? The dream is to produce quality indicators automatically in the production process and have parts of the quality report ready when publishing the statistics. Today each subject matter unit reports on quality on their statistics. Furthermore, we have a centralised system for following and improving quality in the statistics. Based on the European Statistics Code of Practice (CoP) we have quality reviews on several statistics each year. Statistics Norway do also report to the Ministry of Finance on timeliness, punctuality, response burden and response rates. Statistics Norway try to combine administrative and statistical systems, such as risk management, internal control system, Lean thinking, quality reviews and other quality reports, to find the best way to measure quality and contribute to continuous improvements in a consistent and efficient way. Some of the lean projects has established standardised procedures for different processes with checklists step by step. These checklists

may be a base for collecting metadata on quality indicators from each step. In the quality reviews some of the major findings are need for better harmonisation of documentation such as "About the statistics" and harmonisation of metadata. When modernising the production system there will be major changes in the production system. This can give us the opportunity to standardise and harmonise production, metadata, documentation and dissemination.

Keywords: Digitalisation, quality system, quality indicators, quality reviews, standardisation and harmonisation

Title of abstract: Quality reviews of official statistics and the role of the external stakeholders: some initiatives from the NSI of Spain

Agustín Cañada, Asunción Piñán, Luisa Muñoz
National Statistics Institute of Spain (INE)

Nowadays, the national statistical offices are moving a step forward from traditional quality assessment done by the offices themselves, towards a more complete quality system that involves external experts and other stakeholders. User Satisfaction Surveys (USS) have remained so far as the traditional tool; they are the core instrument when assessing quality and relevance of statistics. The INE has regularly conducted (every three years) a series of USSs to qualified users starting in 2007. The last round was completed in 2016 and included a set of developments and extensions, such as weighing the different quality dimensions and then obtaining a global composite satisfaction index; or customization of the survey attending to different categories of users (i.e.: users from the media). Now we would like to move towards a greater and intensive participation of different Stakeholders. In fact, this approach constituted a recommendation in the last Peer Review. The INE, to fulfil this recommendation, has launched several improvement actions:

First, a new procedure to evaluate routinely the statistical operations by the statistical system collegiate bodies, specifically the High Council of Statistics (HCS). It is a collegiate body, composed of representatives from different stakeholders and users (business organizations, trade unions, researchers…). Upon this structure, the INE has drawn a procedure to involve the HCS in the quality assessment of the statistical production, and statistical operation (Labour Cost Index) has been selected for a pilot test.

Second, the INE has started up the organization of seminars with experts (researchers and academia) focused in a specific statistic, within the yearly meetings co-organized with the „Statistics and Operational Research Society".

Alongside a round table session on a specific group of statistics has been organized within those meetings, with the participation of key experts in the selected statistical operation. With these actions, the INE of Spain will set up a quality assessment system including external evaluations to the institution about its statistical products, even for the years between peer reviews. However, the actions previously described do not need set aside specific research on users: as the traditional USS are being supplemented by new assessment procedures linked to the DIGICOM Project where the INE is involved. In this paper, all these initiatives are described, emphasizing the means and tools to incorporate experts' views in the quality management system of the INE.

In this paper, all these initiatives are described, emphasizing the means and tools to incorporate experts' views in the quality management system of the INE.

Keywords: quality reviews, external stakeholders, user satisfaction survey, external experts, collegiate bodies, users engagement

Title of abstract: Assessing the quality of the geographical dimension in macro-economic statistics through mirror data

Tjeerd Jellema, Fausto Pastoris, Carmen Picón-Aguilar
European Central Bank, Frankfurt am Main, Germany

The quality of the geographical breakdown in the balance of payments and related statistics such as international trade in goods, trade in services and FDI statistics can best be assessed by means of comparisons with mirror data, to assess bilateral asymmetries. Such checks form a natural complement to the validation of consistency between balance of payments and the national accounts external account. Although such comparisons are performed regularly, they tend to focus on pairs of countries, and do not provide sufficient guidance to assess which of the countries involved has relatively better data, nor do they provide an indication of the level of quality achieved in specifying the geographical dimension. This paper describes three synthetic indicators developed to provide for a group of countries, and specifically in the context of an economic union external account, an assessment of the quality of the geographical breakdowns by country and the relevance of a country to the aggregate asymmetry of that group of countries. The indicators are applied in the context of euro area and the EU foreign direct investment statistics.

Keywords: Asymmetries, mirror data, balance of payments, foreign direct investment statistics

SESSION 05 – QUALITY MEASUREMENT AND REPORTING II

Title of abstract: Quality reporting in the ESS – State of play and next steps

Javier Alcantara Ortega
Eurostat

Member States have to provide Eurostat with reports on the quality of the data transmitted. Eurostat analyses the reports in order to assess the quality of the data and publishes a summary. Modalities, structure and periodicity of the quality reports depend on the different sectoral legislation. But the general framework for quality reporting in the ESS is standardised: quality criteria and a reporting structure are well defined. Four main pillars sustain Quality Reporting. In first place, the legislation about quality in statistics that goes from Article 12 of Regulation 223/2009 to the different sectoral legislation. Secondly, the standards: ESMS (Euro-SDMX Metadata Structure, user-oriented), ESQRS (ESS Standard for Quality Reports Structure, producer-oriented) and SIMS (Single Integrated Metadata Structure, merging the previous two). The third pillar, the technical implementation across all statistical domains. The fourth pillar, the guidelines and advice, with the ESS Handbook for Quality Reports as flagship publication. The challenges ahead are achieving as much as possible a harmonisation of the standards and guidelines used for Quality Reporting across domains. The ESS Handbook on quality reporting should be the first and most important point of reference for all domains, giving plenty examples of use for different sources of data and statistical domains. More centralised work and coordination in each national statistical authority and in Eurostat is also needed to provide support, advice and training of staff. Legislation concerning quality reporting must be harmonised across domains to the extent possible. Cooperation and interoperability of metadata systems must be enhanced, over domains and functions. Countries should also be strongly encouraged to publish their national quality reports.

Keywords: quality reporting, metadata, quality

Title of abstract: Raising awareness and continuously improving quality in Statistics Denmark

Karin Blix
Statistics Denmark

The quality management system in Statistics Denmark is built on three pillars – quality assurance of documentation of statistics (quality reports), quality audits and our process model. The documentation of statistics is based on the Single Integrated Metadata Structure (SIMS). The central quality unit is responsible for quality assurance of the documentation of statistics which is updated with any new publication of statistical products. The audit process is built up around the European Statistics Code of Practice (CoP) and the Generic Statistical Business Process Model (GSBPM) and is conducted by the quality coordinator and a team of internal and external experts. The subject matter units are confronted with principles 4-15 in the CoP and are asked to fill out a self-assessment form based on these principles. They are encouraged to consult the Quality Assurance Framework (QAF) during the self-assessment. The idea was to think big and start small. The first statistical products were chosen from each of the departments – economic, social and business statistics. In the second round of auditing, statistical products were chosen to involve as many middle managers

as possible and in this way promote CoP in every corner of the organisation. In the third round of auditing we went for central and more complex statistical products also involving external experts in the process. The next step is to involve the users more intensely by conducting interviews in focus groups comprising central users of the statistical product. In the paper the process of the continuous improvement will be described with examples and additional initiatives taken to raise awareness of quality in statistics will be presented.

Keywords: Quality auditing, Quality management, Code of Practice, GSBPM, Continuous improvement

Title of abstract: Coherence – Accuracy – Flexibility

Carsten Schumann
Federal Statistical Office of Germany (Destatis)

In European Business Statistics the "number of active enterprises" and the "number of persons employed" are provided from two sources ("Structural Business Statistics" (SBS) and "Business Demography" (BD)) with equal populations and similar definitions. Consequently, users expect the same results. However, deviations occur because most Member States use sample surveys for SBS and administrative data processed in the Business Register (BR) for BD. In the past some deviations have been severe and Eurostat launched initiatives to remove the gap. Coherence can be improved to a certain degree by calibrating results. However, it is difficult to decide on a benchmark: primary surveys have a sophisticated quality assurance for high reliability of each characteristic while the BR respectively the administrative data has better coverage and allows the publication of small-area data. Unfortunately, the NACE-code – a very crucial characteristic for business statistics – is not always reliable from administrative sources and SBS has to cope with non-responses. Furthermore, there is a limit to the amount of calibration constraints and not calibrated variables and breakdowns can be biased. Clearly, this results in a delicate trade-off between credibility of published results, their accuracy and flexibility in promptly producing additional results on newly arising user demand. Recent initiatives promoted "from stove pipe to data warehouse" and "reuse of existing data" to enhance flexibility and improve responsiveness while reducing response burden. Pilot projects have proven the vast potential of micro-data-linking and are still just tapping the surface. Ex-ante restrictions imposed on results before publication might hamper these possibilities. This presentation wants to trigger a discussion about goals and limitations of initiatives to improve coherence. Should we bend results to enforce coherence? Is coherence needed on every level of aggregation? How much gap can be explained to users? How will calibrations affect accuracy and flexibility?

Keywords: Coherence, Accuracy, Flexibility, Consistency, Credibility, FRIBS

Title of abstract: A modest attempt at measuring and communicating about quality

Laurie Reedman
Statistics Canada

While we were mostly disseminating aggregate statistics derived from a sample survey, we used the sampling error as the backbone of quality reporting. Now we are moving towards disseminating aggregate statistics derived from non-survey methods, and disseminating micro-data products. What should we be communicating about the quality of these products? What do we as data producers need to know about the quality of data is it traverses our processing steps, and what do the users of our data products want or

need to know about its quality? This paper explores these questions in the Statistics Canada context. We are looking for common vocabulary, and a standardised way to represent different aspects / attributes / dimensions of quality. In particular we are exploring ways to measure and report accuracy, and how to reflect the impact on accuracy of processing steps such as data integration, imputation and disclosure avoidance. Finally we look at how to summarize and communicate the accuracy of a data product in a way that informs use.

Keywords: quality, accuracy, quality reporting, quality dimensions

Title of abstract: Fulfilling the ES Code of Practice and assessments of quality in official statistics

Heather Bergdahl, Joakim Malmdin, Ingegerd Jansson
Statistics Sweden

A new self-assessment of the quality in official statistics has been launched for the Swedish System of Official Statistics, the results of which are reported annually to the Swedish government. The objective is to stimulate quality improvements in official statistics and provide a basis for the government to make a "Commitment on Confidence in Statistics". In 2013 the Official Statistics Act in Sweden was amended to include quality criteria comprising relevance, accuracy, timeliness, punctuality, accessibility and clarity, comparability, and coherence (viz. Regulation (EC) No 223/2009, Article 12). In 2016 Statistics Sweden issued executional regulations for the System as a whole specifying an updated quality concept with five main components to be used to describe quality in the development, production and dissemination of official statistics. In 2017, additional regulations were issued in the form of a questionnaire for statistical authorities to annually carry out self-assessments of the quality in their official statistics. The questionnaire is based on the updated quality concept and the focus is on change in product quality over time. An essential starting point in the assessment is the purpose of the statistics and user information needs, which are sub-subcomponents in our updated quality concept. Statistical authorities work actively to strive towards compliance to the ES Code of Practice (CoP). In this context, a most relevant question has been raised as to how our new self-assessments assist us in our efforts towards this goal. Performing assessments is fundamental for making continuous improvements. The CoP refers to assessments in connection to several principles besides principle 12, Accuracy. In this paper, we will explore how approach to assess quality in official statistics can be a useful and potentially powerful tool for continuous improvements and a means to achieve quality improvements in official statistics as well as greater compliance to the CoP.

Keywords: Assessment, quality concept, continuous improvements

SESSION 06 – COORDINATION OF NSS, ESS AND BEYOND

Title of abstract: Results of intensive development actions in official statistics coordination – a status report from Hungary

József Kárpáti
Hungarian Central Statistical Office, HCSO

The paper/presentation will introduce how the Hungarian Central Statistical Office developed new actions and procedures in order to enhance its coordinative function. In the first chapter it presents the coordination tools introduced by the new national statistics act which entered into force on the 1st January, 2017. The next chapter introduces the National Statistics Code of Practice in detail. This is a new set of guidelines which were created to foster and uniform the performance of this public duty on a high quality level in Hungary. The present members of the Official Statistical Service committed themselves to the guidelines laid down in the Code, and also go under an audit procedure during 2018, where their compliance is examined. The third part introduces the methodology of these audits. The fourth part of the paper/presentation will summarize the experience and impressions gained during the already implemented audits of the organisations. The next chapter describes the results of the newly established National Statistics Coordination Board, primarily focusing on how the Board will be involved in the monitoring of development actions, related to the National Statistics Code of Practice audits. The final part of the paper/presentation would describe the further coordination initiatives of the Hungarian Central Statistical Office.

Keywords: co-ordination, statistics act, organisational development, code of practice, quality audits

Title of abstract: Improving the quality of national statistical systems

Vincenzo Lo Moro
Italian National Institute of Statistics (ISTAT)

The success of a Modernization project in an NSI depends on the validity of the project as well as the quality of implementation. Usually the project is carried out within the same organization, the NSI. Problems arise when the scope is the Modernization of a National Statistical System. Aside from the coordination role of the NSI, each organization where there is a statistical service (Ministries, Regions, Municipalities) has its own logic, mechanisms and hierarchy. The Italian Statistical System (SISTAN) is both broad and incisive. To maintain an acceptable level of quality the instruments used are somewhat different than those used within ISTAT. The overriding characteristic of these instruments is that they are "soft tools", not coercive, but inclusive and functional on a cultural level. In essence, results are achieved in cooperation, a "grey area", not as the mandate of a single organization. ISTAT has taken a number of initiatives towards improving the SISTAN Statistical Offices performance. From 2014-2015 ISTAT underwent the Eurostat peer review. This was the backdrop to the program which was launched in 2016-2017. This program has three main initiatives: the improvement in statistical production by issuing Quality Guidelines tailored for the SISTAN producers, and subsequent statistical audit of the most relevant processes. The update of the list of the Other National Authorities (ONAs) based on common criteria, which should be coherently applied by Member States. An intensive high-level training (75 hours in 15 days) for central Statistical Offices staff based on the Italian Code of Practice for Statistical Quality. The paper presents these three initiatives and what has been learned from the experience (results obtained anddifficulties). It will conclude with a preview of upcoming programs.

Keywords: National Statistical System, Other National Authorities, Modernization, Organization

Title of abstract: Official statistics in Denmark

Peter Bohnstedt Anan Hansen, Kim Voldby Pedersen
Statistics Denmark

With Regulation 223/2009 on European statistics a legal requirement to coordinate European statistics is introduced. The heads of the National Statistical Institutes of the Member States are now responsible for producing national quality guidelines and for monitoring their implementation in order to ensure compliance with the European Statistics Code of Practice (CoP) in all nationally produced European statistics – including European statistics produced by the so-called Other National Authorities (ONAs). In parallel with the implementation of Regulation 223/2009 Statistics Denmark is working on the introduction of the concept of "official statistics", covering not only European statistics but all statistics produced by public authorities in Denmark. In this regard the requirement of monitoring the ONA's compliance with the CoP is also used in a broader national context as a lever to ensure official statistics of a commonly recognized quality standard – a national quality stamp for statistics. For the purpose of monitoring the ONA's compliance with the CoP for European statistics, Statistics Denmark has developed a focused set of guidelines based on the CoP in order to make it practically applicable for authorities that do not have statistical production as their main task. The ONAs will annually submit a self-assessment to Statistics Denmark on the compliance. Together with a "mini peer-review" this forms the basis for a report to Eurostat and to the Danish Government on the compliance with the CoP. For the purpose of labeling not only European statistics but all statistics produced by public au-

thorities as "official statistics", the national producers of statistics are encouraged to ensure compliance with the principles of the guidelines. All public produced statistics that comply with the guidelines developed on the basis of the CoP will be labelled "official statistic", ensuring a common standard for official statistics at European and national level.

Keywords: Official statistics, National quality guidelines

Title of abstract: Coordination of National Statistical System – Experience from the Hellenic Statistical System (ELSS)

Christina Pierrakou - Hellenic Statistical Authority (ELSTAT), Stamatios Theocharis - Statistical Head and Head of Internal Audit Unit of Ministry of Interior, Greece

In the recent revision of the European Statistical Code of Practice (CoP) a new Principle for "Coordination and Cooperation" is to be added, in line with the amended Regulation 223/2009. The new Principle consists another step towards the enhancement of cooperation in National Statistical Systems. The National Statistical Systems (NSSs) have not the same degree of centralization. Consequently, there is a challenge faced by most National Statistical Institutes (NSIs) about the amount of European Statistics produced by Other National Authorities (ONAs). Thus, it is essential for NSIs to have a strong position in the NSSs stipulated in the national statistical legislation and recognized by the ONAs. In this paper, the experience from the coordination role of ELSTAT in the Hellenic Statistical System, which has a certain degree of decentralization, is described. Furthermore, the ONA's experience from the certification process of "official" statistics is presented. The key role of the Statistical Head in the cooperation mechanism within the ELSS for the implementation and maintenance of high quality standards in the development, production and dissemination of statistics by ONAs is also described. Opportunities for better cooperation at national level as well as the further consolidation of citizens' confidence in the production processes of statistical products, are analyzed through the case study of the Ministry of Interior, which has recently established a relevant organizational structure for the Statistical Head.

Keywords: Coordination, certification process of official statistics, ONA, Hellenic Statistical System

Title of abstract: Common Quality Reporting System for Official Statistics of Finland

Outi Ahti-Miettinen, Kari Djerf
Statistics Finland

Statistics Finland, in cooperation with Finnish Customs and Natural Resources Institute Finland, designed and built a common quality reporting system for all organizations producing Official Statistics of Finland. Furthermore, the objective is to unify the structure and quality of the quality reports in all the official statistics of Finland. System was developed for collecting, storing and disseminating quality reporting using Single Integrated Metadata Structure (SIMS) v 2.0 endorsed by the ESSC in November 2015. System was constructed using Service Oriented Architecture and all producers of official statistics will be invited to use the system through internet interface.

We will present main contents and general technical solutions of the new quality reporting system as well as its properties and links to general dissemination portal.

Keywords: Quality reporting system, Official Statistics, SIMS v 2.0, Coordination, Cooperation

SESSION 07 – EFFECTIVE PARTNERSHIPS FOR QUALITY IMPROVEMENTS

Title of abstract: Legal aspects related to enhance the cooperation with data holders: The case of Spain

Gregorio Izquierdo Llanes, Antonio Salcedo Galiano, Yolanda Gómez Menchón
National Statistics Institute of Spain (INE)

The concept of cooperation is closely related to coordination in the Regulation 223/2009 and in the Code of Practice of European Statistics. Cooperation is understood as an horizontal tool between national statistical authorities and other public entities or between statistical authorities and the scientific community, or even among statistical authorities of different countries (international cooperation), but the ESS Quality Declaration adopted in 2016 introduces a new meaning of cooperation. The first question that comes to mind is if the ESS is moving from a traditional cooperation among public data providers -the owners of the administrative data- to an enlarged cooperation adding the private sector. Addressing a very hot issue like big-data, in this article we analyse some issues such as the need for legal support for cooperation with data holders or which type of legal acts could be more useful for the statistical community.

Keywords: European Statistics, ESS Quality Declaration, Big Data, Cooperation, institutional environment, legal framework

Title of abstract: Ethical implications of using Big Data for Official Statistics

Albrecht Wirthmann, Fernando Reis
Eurostat

We are currently experiencing an all-embracing digitalisation of our societies and economies. The pervasive nature of information and communication technologies is leading to the ‚datafication' of most of our activities and relationships. This development is producing so far unknown amounts of data. New developments in Information Technologies do not only allow capturing but also storing, linking and analysing these data to infer conclusions on data subjects in massive amounts. This new ability might impact persons' everyday life in different ways, positively or negatively. Ethical implications thereof have already been discussed in literature. The proposed paper will discuss the implications of using big data in official statistics. These implications will differ from risks and hazards related to big data usage in general. Firstly, Statistical offices are not targeting individual subjects but only aggregated results and identified patterns. Secondly, there are already principles and ethical guidelines concerning the statistical confidentiality of personal data. Therefore the paper will focus on issues of data quality, availability of data sources, dependence on third party sources, data manipulation influencing the results, transparency of data and methods of data analytics, or scientific approach to data analytics. The paper will identify issues and analyse possible consequences for official statistics and contrast them with the existing ethical frameworks (Fundamental principles of official statistics, European Statistics code of practice, Declaration on Professional Ethics). Finally the paper will define or emphasize relevant principles on the ethics of using big data that could be followed in the process of integrating these new data sources into Official Statistics.

Keywords: Big Data, Ethics, Code of practice

Title of abstract: Innovation in collection: strategies to improve and maintain the quality of administrative data

Hélène Bérard, Josée Cellard, Cory Lusk
Statistics Canada

Statistics Canada actively obtains administrative data from the public sector, the private sector and various organizations to support the production of official statistics. A number of strategies are in place to facilitate data access and to ensure that the data obtained from various organizations meet specific quality requirements. The strategies implemented to meet these goals include: 1) outreach activities prior to obtaining the data to explain the usefulness of the data, 2) offering technical support, 3) organizing data quality workshops to promote the use of sound methods, 4) identifying specific requirements in data acquisition agreements to ensure the quality and timeliness of the data, 5) negotiating access to a test file to assess fitness for use before proceeding with an official acquisition, 6) establishing a Quality Evaluation Framework to ensure consistency and completeness of assessments across the departments many administrative data files and, 7) maintaining post acquisitions communications, and 8) developing mutually beneficial relationships to ensure long term supply, usefulness and quality of the administrative data. One of our current challenges is to identify and influence future changes to administrative data sources that may affect statistical use. This paper will discuss specific examples to illustrate the challenges faced and the lessons learned in establishing these various strategies.

Keywords: administrative data acquisition, partnerships, quality assessment of administrative data

Title of abstract: Access to Big Data for statistical purposes

Thomas Riede, Natalie Rosenski
Federal Statistical Office of Germany

The topic big data is of high importance for all statistical offices as well as the Federal Statistical Office of Germany (Destatis). On this account, Destatis is a member of the European Statistical System (ESS) Task Force, the ESS Steering Group as well as the ESSnet Big Data. Furthermore, Destatis is a member of the United Nations Global Working Group on Big Data for Official Statistics. Destatis has already started several feasibility studies concerning big data, e.g. web scraping for price and labour market statistics. Furthermore, Destatis has entered cooperations with T-Systems, which is a subsidiary of the Deutsche Telekom AG, and the German Aerospace Center to use mobile phone and satellite data for the examination of its usability in official statistics. To be able to use big data for official statistics and not only for feasibility studies, individual cooperations with private enterprises are not sufficient. Instead, a permanent access to data held by private enterprises is necessary, which is unattached of their willing or the market situation. Therefore, Destatis is aiming for a legal basis to get access to privately held data, on which further cooperations can be premised on. Big data should then be free of charge, but the service of the enterprises to prepare the data for statistical purposes could still be charged by the enterprises. Fees for this service enable the access to the knowledge of the data providers. Furthermore, official statistics could get semi-final products by the data providers in order to save money within the offices. A benchmark for this legal foundation is delivered by France and the United Kingdom.

Keywords: Big Data, data access

Title of abstract: Big Data Strategies for Official Statistics

Peter Struijs, Sofie De Broe
Statistics Netherlands

The environment of National Statistical Institutes (NSIs) is rapidly changing in many respects. The emergence of new data sources provides a number of opportunities for official statistics. At the same time this creates challenges, including how to deal with quality issues when for example developing so-called experimental statistics and turning them into official statistics. As many NSIs have started using big data for statistics, the need for a strategic approach has become increasingly clear. The paper describes and assesses strategic options and explains the big data strategy of Statistics Netherlands, which is now being implemented in the Dutch Center for Big Data Statistics. In essence, big data strategies are about positioning NSIs in the changing environment. The paper identifies the true game-changers for official statistics and formulates the associated strategic questions. Their answers depend on where the value added of official statistics is sought, which is to no small extent related to quality considerations associated with the use of new data sources. New approaches may be called for. In any event, the role of NSIs is bound to change. The traditional role of quasi monopolistic provider of statistics on the many facets of society will erode through the rise of competition. However, the institutional and professional foundation of NSIs may also be exploited for assuming new roles. Ideally this will result in a society that is better informed about relevant phenomena and better equipped to counter tendencies where the value of facts is discredited.

Keywords: big data strategy, big data vision, new data sources

SESSION 08 – HUMAN RESOURCES AND QUALITY CULTURE

Title of abstract: Capabilities – Improving the quality of statistical capacity development

Steven Vale
UNECE, Switzerland

Statistical capacity development is a growing priority for many national and international statistical organisations. They are faced with the challenges of increasing demands to produce new types of statistics, to use new data sources, and to master new technologies. The growing data requirements for monitoring progress towards the sustainable development goals, as well as to provide evidence to address socio-economic challenges such as globalisation and migration, all add to the pressure to increase efficiency. Increasing the quality and effectiveness of statistical capacity development activities is essential if statistical organisations are to meet these challenges. There is a growing recognition that capacity development should be about much more than just providing training courses for staff. A more holistic approach is needed, focusing on improving the capacity of the organisation as well as the capacity of the people who work for it. This paper explores how new tools and concepts, adapted from those used outside the official statistics community, can help. It focuses particularly on the concepts of "capabilities" and "maturity models", and considers how these concepts can be used to identify priorities more efficiently, and to monitor the effectiveness of capacity development activities.

Keywords: capacity development, efficiency, capabilities, maturity models

Title of abstract: A matrix model for human resource organization to improve effectivness and efficiency in official statistics – ISTAT

Silvia Losco, Sabrina Pifferi, Manuela Cola
Italian National Institute of Statistics (ISTAT)

Quality does not only mean quality of products, services, processes and methodologies, but quality is, first of all, an approach that invests all the organization. In 2016, the Italian national Institute of Statistics (Istat) launched the Modernisation Programme, whose main objective is to enrich the supply and quality of the information produced, while improving the effectiveness and efficiency of overall activity. For what concerns organisational items, Modernisation Programme based on two main focuses regarding centralization of corporate support services, introduction of a portfolio and project management increasing the attention on statistical outputs and the efficient management of resources. In this frameworks, Istat switches from a functional organization to a matrix model of organization that aims to achieve a better human resource management, crucial in a diminishing resources perspective. In Istat matix model, staff knowledge and skills can be shared between functional departments and project teams according to needs. In this organizational model, people who work on projects have basically two leaders: the authority of the functional manager runs vertically downwards and the authority of the project manager runs horizontally. Precisely this crossing between the reporting lines determines the meaning of the matrix. Istat has been experimenting a matrix organizational structure for almost 2 years and we have now elements for an initial assessment of the model. The paper describes how this model has been introduced at Istat, the advantages – both real and potential – of this structure in a statistical environment and the problems to cope with for a full implementation.

Keywords: matrix model organization, human resource management, modernisation, ppm

Title of abstract: Sensibilization on quality at the training center of Insee

Xavier Helfenstein
INSEE - Cefil, France

In the training center of Insee (Cefil), the culture of quality is disseminated to the futur staff of the NSI through a process of acculturation. Quality is omnipresent in all the courses and is presented to the trainees in various dimensions going from methodological to ethical dimension.

Different principles of quality are approached through study projects. For example, the project of concerted analysis of statistical tables demonstrates the necessity for the statistician to present the results of his works in a clear and understandable way. By organizing a forum, the skill transmission project emphasizes the importance of pedagogy. The statistical survey project places the trainees in situation to build a collection of data from scratch in order to answer to the request of a public actor. This exercise requires to use every skills of a collective, to meet the deadlines, to restore results in compliance with the statistical secret and to document the data. It teaches the trainees how to manage the impact of a data processing on the final result. The Cefil also offers a classic sequence on quality to the trainees but the education of the training center wouldn't be efficient enough to the objective of professionalization if it dispensed only this sequence. Indeed, the support of quality is observed to be more induced by a daily behaviour than by a knowledge of an academic subject. Operating in project mode takes there all its interest. The approach of quality is thereby embedded in a holistic conception of the educational route of whom it is a component.

Keywords: Quality culture, training, education, learning by doing

Title of abstract: Matching pairs, agile and pin the tail on the donkey – making quality management relevant to non-statistical staff

Victoria Staples, Orlaith Fraser, Olivia Wayman, Ian Tucker
Office for National Statistics, United Kingdom

Embedding a culture of quality is a key goal for those working towards the 2021 Census, but many non-statistical staff struggle to understand how quality relates to their roles. We gathered feedback from colleagues representing a range of professions to understand these issues and redesigned an introductory quality management course to tailored to the specific needs of the Programme. The course focuses on three key learning objectives; what quality is, why it is important and how we will achieve good quality outcomes, with the aim of broadening its relevance to, and better engaging staff to achieve a quality culture. We started with our corporate quality management strategy and framework, extending this to include additional components for those working in programme management, procurement and technological areas. We then considered how to increase engagement with our course content. Using best practice from other successful trainers, we adopted a more informal approach while maximising impact through collating real-world examples from across the 2011 Census operational experience and a range of professions. We also highlighted the relevance of quality in day-to-day activities through linking the components of our quality framework to existing Agile working practices, demonstrating that staff were already 'doing' quality management whether they realised or not. We introduced simple group activities to raise awareness of the dimensions of statistical quality, our programme quality objectives and activities relevant to attendee's own roles. Linking these examples to the components of our framework and the course objectives and the use of online quiz-

zes reinforced the key learning points. The course has received outstanding feedback for its interactive and inclusive approach which has enabled staff to better engage with its content and understand the relevance to their own roles. This has translated to increased understanding of quality and a higher profile across the programme.

Keywords: Quality management, training, increasing engagement, quality culture, agile

Title of abstract: The new quality strategy in the modernised Italian National Statistical Institute

Giovanna Brancato, Giorgia Simeoni, Andrea Bruni, Gabriele Ascari
Italian National Institute of Statistics (ISTAT)

Under the pressure of common drivers for innovation, the National Statistical Institutes of several countries, including the Italian institute (Istat), have implemented modernisation programmes in their organisations. They are broadly aimed at improving the amount and quality of the statistical information produced, while increasing the efficiency and cost-effectiveness of the organisation. Quality management has a prominent role in the modernisation programmes being an overarching process supporting the organisation at institutional, process and product level. Starting from the early '90s, the long-standing Istat quality management investment has produced an advanced and consolidated quality strategy, harmonised with Eurostat guidance, whose main pillars rely on: the setting of shared standards; a spread network of quality pilots; an articulated assessment approach aimed at continuous quality improvement. The initiatives developed so far at Istat, although generally valid, need to be further improved and tailored with respect to the changed statistical production environment and to the shift from an organisation based on domain specific-silos to a more integrated production model. In addition, the tools already developed for documenting, measuring and assessing quality of statistics based on surveys and/or administrative data need to: I) be generalised to cover the unstructured data sources; II) be managed in a different organisational setting and III) be extended to consider process performance. The paper will present a proposal on how to re-organise the quality activities to better support the needs of a modernised statistical organisation. Firstly, each element of the quality strategy will be analysed in the light of the modernised Institute and of the recent orientation provided by the European Statistical System Common Quality Framework. Then, a redesigned approach to process and product quality assessment and improvement will be proposed. Finally, the recent activities focused on the quality improvement of the National Statistical System will be shortly described.

Keywords: modernisation of official statistics, quality framework, quality strategy

Title of abstract: Future – Engage – Deliver: refreshing the UK Code of Practice for Statistics

Ed Humpherson
UK Office for Statistics Regulation

We live in a data-rich world that is rapidly changing. This environment impacts us all, as well as the very statistics used to make essential judgements and decisions in all parts of society. In the UK the Code of Practice for Official Statistics has been at the heart of setting high standards since 2009. The changing data landscape and recommendations from Sir Charles Bean's Independent Review of UK Economic Statistics provided a strong impetus for considering how to ensure the Code continues to remain relevant in driving further improvements in official statistics. We followed a Future – Engage – Deliver model. We envisioned the Future in our Code stock take, conducted from late 2015 to December 2016. Our testing and development of the refreshed Code in the period January to June 2017, in which we conducted a series of focus groups, as well as a formal consultation between July and October 2017, represented the Engage phase. And our Code Consultation informed our Deliver phase, in which we have revised the Code content to reflect stakeholder feedback. The three pillars of Trustworthiness, Quality and Value now form the central framework for the statistical practices. We are focusing on embedding the Code through the development of interactivity and guidance through our website. And we are advocating the application of the Code beyond official statistics, to have a wider reach. Our ambition is for organisations publishing statistics outside of the official statistics community to also apply the Code principles. The fundamental components which form the basis of the Code can acts as guidelines for all publishers of statistical information for the wider benefit of society.

Keywords: code of practice, statistical authority, trustworthiness, quality, value, regulation

SESSION 09 – USE OF GEOSPATIAL INFORMATION TO INCREASE QUALITY

Title of abstract: How geospatial information adds value to existing sub-national data and territorial typologies

Valeriya Angelova Tosheva, Hannes Reuter
Eurostat

The mission of Eurostat is to provide high-quality statistics for Europe to e.g. support policy decisions. In line with a service oriented offering the regional statistics and geographical information team is providing methodological support to policy-makers, researchers and the general public to better understand the existing data but also offering new ways of meeting data requirements in an innovative and flexible manner. Within Eurostat, GISCO is responsible for meeting the European Commission's geographical information needs. Statistics in GIS environment makes it possible to process, store, analyse, aggregate and integrate different data sources with the final aim to visualize and support policy decisions. Additional, it allows to produce unique sub-national indicators that are out of the scope of the existing at Eurostat and the National Statistical Institutes data collections. Thereby, GIS enables a transformation of data available by official territorial typologies into data by a customised territorial type. This paper describes two recent examples from the Eurostat's everyday practice of how geospatial information has been used in order to satisfy data requests. We show how for Directorate General for Agriculture and Rural Development (DG AGRI) the number of schools, kindergartens, post offices, sport clubs and bars as social meeting points in individual rural regions per 100 000 inhabitants have been derived. Second, how for the Directorate General for Maritime Affairs and Fisheries (DG MARE) various demographics and socio-economic data by maritime ports in Europe have been derived. Geospatial information proved essential to support policy decisions in the impact assessment of the Common Agricultural Policy domain as well as providing reliable statistical estimates for the blue growth.

Keywords: GIS, sub-national data, user needs, territorial typologies, reengineering existing data, alternative data sources, service orientation

Title of abstract: Communictation as a Quality Factor in Multi-source Spatial Data Integration Process

Włodzimierz Okrasa
Statistics Poland and University of Cardinal Stefan Wyszynski in Warsaw, Poland

The two-fold goal of the proposed presentation - Communication as a Quality Factor in Multi-source Spatial Data Integration Process - is (first) to explore and systematize the quality-related issues arising specifically in the context of work with spatial / geo-referenced data, and (second) to recognize and discuss the role of communication within the statistical process and, subsequently, in constructing an analytical (multi-source) database through integration of data from different sources. The main focus will be on integrating survey and administrative data for the purpose of multi-level spatial modeling of community well-being and subjective (individual) well-being, toward a general methodological framework of ‚spatially integrated statistical research'. With ‚horizontal' - across space - and ‚vertical' - across different types of units, like persons

or households and local communities -dimensions of producing and linking/matching the geo-referenced data, taking an evaluation-oriented policy analysis perspective. Communication is then understood both as an intra-statistical and between data producers and data users occurrence, including methodological (error -reduction) and institutional (e. g. ‚statistical knowledge' communication) procedures, remaining basically under responsibility of the official statistics. Since modeling relationships (interaction and influence) between community well-being and individual (household) well-being provides most of the challenges associated with multi-level analysis (modeling) of spatially dependent phenomena, this problem will constitute the main part of the empirical demonstration of the proposed, better communication for better quality' approach.

Keywords: Communictation, statistical process, spatial data integration, multi-level modeling

Title of abstract: GEOSTAT 3 – A European Statistical Geospatial Framework

Karin Hedeklint
Statistics Sweden

Members of the ESS (European Statistical System) are currently seeking to collectively modernise their statistical production systems, to transform their operations and to derive new relevant metrics and indicators. This modernisation process involves an ambition to level up integration of geospatial information in statistical production. One of the main goals of the GEOSTAT 3 project is to develop a European Statistical Geospatial Framework (ESGF), building on the principles of the Global Statistical Geospatial Framework (GSGF). The development of an ESGF is believed to be an important step towards a better and more coordinated integration of statistical and geospatial information. The ultimate objective of the ESGF is to increase resource efficiency, to obtain a higher degree of harmonisation between countries and to support creation of a more flexible statistical production. Whereas the Global Statistical Geospatial Framework is a high-level framework consisting of five generic principles broad enough to apply to any region around the globe, the ESGF aims to be a tighter framework reflecting the specific European situation. It will build on the major achievement already been done for the availability of geospatial information through the INSPIRE directive and the Spatial Data Infrastructures set up by Member States. The GEOSTAT 3 project runs until the end of 2018, but for the 2018 Quality conference, a presentation will be made on the key-elements of the framework as proposed by the project.

Keywords: Geospatial, Framework, Data Integration, Infrastructure, Accessible

SESSION 10 – QUALITY AS A MULTI-DIMENSIONAL CONCEPT

Title of abstract: Quick statistics – how to deal with quality?

Pertti Taskinen
Statistics Finland

Labour Force Survey needs a reservoir of information which is collected directly from the interviewees. Telephone interviews are often quick and even hasty. In this presentation, the object is to tell what kind of quality checks are made to ensure the validity of the data which forms the official monthly unemployment rate - a quick and an important statistics. Mostly, checks are mode independent, but the increasing use of the web data collection mode will be taken account in this presentation. During the data collection stage, the best guarantee of the quality is a trained interviewer who uses the data collection software with certain question-specific rules or instructions. The instructions of the interview should be so simple as possible and to be adapted as commonly as possible. For the Labour Force Survey, it is also recommendable to use the data from the previous round to the next one. This makes an interview easier for both sides - an interviewer and interviewee - without to endanger the quality. At the end of each survey month, the data production model uses three separate programmed checks: one for the collected data, one for the domestic variables and one for the EU variables. The software at the data production has been made user-frendly. If there are no unexpected problems e.g. with databases, the software program is possible to execute over a day. Basically, the figures are then ready but during the dissemination step some working days are needed to draft the publication and translations, and to update databases. In this presentation, the data production model of the Finnish Labour Force Survey is also examined: could we do something even quicker and possibly in a little less resource demanding way with the existing quality? And eventually, what is the best practice to recommend?

Keywords: Quality check, Labour Force Survey, data collection, production model

Title of abstract: The impact of the consistency debate on increased accuracy and coordination within the national accounts and between the national accounts and balance of payments statistics of the EU

Robert Obrzut, Francesca Tartamella
Eurostat

The quality assurance framework of the European Statistical System suggeststhe critical assessment of data sources, statistical techniques and revision practices as well as the assessment and validation of intermediate data and statistical outputs. They point at underlying statistical compilation processes that involve the use of primary data sources and statistics in order to obtain a finalised statistical product. As a prominent example in macroeconomic statistics, the national accounts and balance of payments statistics complement each other in such a manner. This sequential concept, however, hardly reflects the realities of statistical compilation practices, where statistical products are often released in parallel, playing a role both as data source and final statistical product. It is argued that statistical compilation rather has to be perceived as a twinning process which is usually conducted by more than one compiling institution and obliges compilers to introduce only data sources and estimation practices which they can directly control for the sake of

timeliness of their statistical products. In such situations a critical assessment as demanded by the international framework is conducted by each compiler autonomously. In the EU this has resulted in about half of its Member States still releasing national accounts and balance of payments statistics with either high or moderate inconsistencies due to the autonomous use of data sources and compilation practices, although the methodological standards would require full consistency. Different deadlines in the transmission program and limited resources of national statistical institutes can lead to inconsistencies across tables also within the same domain (national accounts), even when the compilation is carried out by the same institution. This paper will combine data evidence about such inconsistencies and concludes that there is a distinct need for increased coordination of the underlying compilation processes both at national and international level.

Keywords: national accounts; balance of payments; trade asymmetries, cross-domain consistency; international comparability; accuracy and timeliness

Title of abstract: Measuring the quality of commercial and big data sources for official statistics

Kimberley Brett, Frantisek Hajnovic
Office for National Statistics, United Kingdom

There has been an increasing shift towards using data science techniques and accessing alternative big data sources across all sectors of society. The UK government has been exploring the use of big data within official statistics. These are neither administrative data, nor existing government data, and could include; web scraped data, big data from commercial companies and social media data. The Office for National Statistics (ONS) is exploring the use of these alternative data sources as part of its „Better Statistics, Better Decisions" strategy. Using these data sources and integrating them into official statistics presents significant methodological challenges, including bias, variety of data (such as text and images) and that there is no control over data supply. As an official statistics producer, ONS is committed to ensure that any statistics derived from this data meet user needs and are of high quality standards. As an emerging field, there is little guidance on the measurement of quality within big data and data science applications for official statistics. Nevertheless, there are existing dimensions of quality that offer a good framework for using data in an official capacity. Derived from the European Statistical System dimensions of quality, the UK Government Statistical Service's (GSS) eight quality dimensions for all published statistics can be applied to big data. These include; relevance, accuracy, timeliness and coherence. The presentation will explore the quality implications of big data and data science methods through example projects of the ONS Big Data team, including use of social media data and collaborations with other National Statistical Institutes. Through these, we will establish how quality could and has been measured using the eight dimensions, challenges faced and ideas for applying quality measures to future big data sources.

Keywords: Data Science, Big Data, Official Statistics, Quality

Title of abstract: Further Challenges in Quality Management of Statistics at the Bank of Japan

Kuniko Moriya
Bank of Japan

The Bank of Japan compiles various economic statistics, which draw a great deal of attention from media, economists, and policy makers. These statistics include financial statistics, price indexes, and the Tankan (a short-term economic survey of enterprises in Japan). As the central bank of Japan, the Bank faces two challenges in compiling statistics: it relies solely on the voluntary cooperation of data contributors when collecting the data, while it also needs to meet high expectations from the public for good quality management of its statistics under high labor constraints and budgetary limitations. The Bank ensures the quality of its statistics by making them comply with "The Basic Principles for the Compilation, Release, and Development of Statistics," published by the Bank in 2009. The Bank also makes continuous efforts to further improve the quality management of statistics. This paper considers recent challenges facing the Tankan, and the Bank's efforts toward enhancing data quality -- including visualization and dissemination -- in order to engage users and meet their demands in a cost-efficient and responsive manner. Specifically, the paper explains the motivation behind these efforts, the process of study, and the desired results. The paper also considers quality management frameworks such as the European Statistics Code of Practice developed by Eurostat, and issues related to the European Statistical System's Vision 2020.

Keywords: quality management of statistics, data quality improvement, dissemination

Title of abstract: Cognitive interviewing in the Disability Pilot Survey in Spain

Rogelio Pujol Rodriguez, Maria Angeles Lora-Tamayo D'Ocon
National Statistics Institute of Spain (INE)

Cognitive interviewing is a technique that helps ensure that a survey questions successfully capture the scientific aims by identifying potential measurement error and question wording problems (overall understanding and interpretation of the question, convenience of the examples given, temporary references, type of scoring at an item level, etc.). This method has proven to increase the quality of the collected data. Hence we conducted several cognitive interviews on two questionnaire models prior to the fieldwork of the Disability Pilot Survey. The aim was to analyze the comprehension and quality of the questions and correctly collect those homes where people with disabilities live, and their limitations. For this purpose voluntary subjects with specific characteristics of interest were recruited and interviewed in a laboratory environment. Some of them were asked to respond to a previously assigned questionnaire model via web (CAWI: Computer-assisted web interviewing) while others were telephoned (CATI: Computer-assisted telephone interviewing). All of them were asked afterwards about their responses and feelings while answering the questionnaire. As a result of these interviews, most of the participants preferred the second model questionnaire (the one with multiple response format), several questions changed their wording (the known GALI question was rephrased) and we confirmed the importance of the examples given at the end of each one of the specific disability question.

Keywords: Pretest, Data Collection, Survey, Disability, Health

SESSION 11 – QUALITY OF MULTI-SOURCE STATISTICS

Title of abstract: Confidence intervals for register-based statistics?

Alexander Kowarik, Johannes Gussenbauer, Eliane Schwerer, Christoph Waldner
Statistics Austria

Most of the time register-based statistics are published without quantitative measurements for the estimation error, e.g. a confidence interval. Based on an indicator proposed in the European project 'Quality of multi-source statistics' (KOMUSO), a resampling method is applied to the Austrian register-based labour market statistics with the goal to quantify the model error of the imputation process and the measurement error of the target variable in the administrative data. A survey data set is used to benchmark this measurement error and estimate transition probabilities to different states in the categorical target variable. A resample method is then applied to generate repeated estimates which are used to compute confidence intervals.

Keywords: confidence interval, bootstrapping, register data

Title of abstract: Quality evaluation of registered-based research

Bart Bakker
Statistics Netherlands and VU University Amsterdam

Due to the inclined availability of administrative registers, more and more register-based statistics are published. Recently developed quality frameworks have tried to mimic the Total Survey Error approach for register data. Main indicators are under and over coverage, measurement error and linkage error. However, these quality frameworks do not present the methods to measure these indicators. In this paper, we present capture-recapture methods (CRC) to estimate under coverage of already linked registers. Over coverage can be determined and corrected for by removing records of people that do not belong to the target population. Duplicates can be identified by linking the records in the combined registers to each other. Measurement error can be estimated by Structural Equation Models (SEM, for numerical variables) and Latent Class Analysis (LCA, for categorical variables) with a measurement component if another source that measures the same concept is linked to the register data. Linkage error can be estimated using probabilistic linkage methods. However, none of these methods is error free in itself. Linkage error also could have impact on the outcomes of CRC, SEM and LCA. The paper show the interdependencies between these methods.

Keywords: register data, under coverage, measurement error, linkage error

Title of abstract: Quality evaluation of statistical processes based on administrative data: a new version of the TSE approach

Orietta Luzi, Fabiana Rocci, Roberta Varriale
Italian National Institute of Statistics (ISTAT)

Over the last decade, National Statistical Institutes (NSIs) have progressively moved from single- to multi--source statistics. By combining different data sources (direct survey, administrative and Big data) NSIs can increase the detail of information, save data production costs and reduce burden on respondents. The Italian NSI (Istat) has strongly increased the use of administrative archives as primary source for statistical production purposes. To this aim, a system of statistical registers based on the integrated use of administrative sources is under development, and many statistical processes have being accordingly re-designed. Such a change calls for a tailoring of the current approaches for quality measurement and assessment. While in Istat a total quality framework based on the Total Survey Error (TSE) is well developed for surveys, a quality framework supporting the design of the new required statistical processes, based on the use of several types of sources, their evaluation and monitoring is still missing. To this extent, the adaptation of the TSE lately proposed in literature for statistical processes using administrative data sources has been taken as reference. In this paper we illustrate as the proposed quality framework has been tested on a new process - the statistical register Frame-SBS that supports the Structural Estimates on businesses- representing a milestone in the new system of statistical registers in Istat. As a major result, the paper contains a proposal for an additional quality assessment phase. It is necessary, indeed, to define an evaluation phase during which each single administrative source is evaluated on its own with respect to the specific statistical aims of the process. This phase should help the main decisions about how to integrate data, that is in identifying the main steps about how to combine the external sources and to outline the specific steps delivering the final output.

Keywords: Administrative data, Quality evaluation, Total survey error

SESSION 12 – PROMOTING THE VALUE OF OFFICIAL STATISTICS

Title of abstract: "An outgoing expert personality" – using a brand platform as a tool to improve external communication

Karin Hansson, Lisa Thiel
Statistics Sweden

Regular surveys show that the general public has great confidence in Statistics Sweden, but the same studies also indicate that the majority know very little about what we do. We also know that many people confuse us with other survey companies, including commercial ones. Furthermore, there seems to be a gap between how we feel about our users and how they see us. We want to be helpful, friendly and engaging but often come across as introvert and hard to understand. To improve our external communication and engage better with users, the communication department in collaboration with HR have created a brand platform. The platform consists of, among other things, our logo, graphical profile, guidelines for using pictures and colours, as well as a personality for Statistics Sweden: „Trustworthy with the ability and commitment to attract people to want to contribute, use facts and wanting to know more."

Statistics Sweden is an outgoing expert personality that is passionate about their job

instills great confidence is contemporary in expression participates in the public discourse. The personality describes how we want the world to perceive us when they visit our website, attend one of our seminars, read emails and letters or report data in our surveys. In short, in any kind of interaction with us. In this presentation we will focus on how we have used the brand platform to communicate better via our Statistics Service, where we answer questions about statistics by phone and email. The most important tool as a series of workshops resulting in their own guidelines on how to express Statistics Sweden's personality linguistically. Through regular audits of their email replies, we make sure the work has led to concrete improvements.

Keywords: branding, brand platform, external communication, statistics users

Title of abstract: Development of activities based on customer experiences and feedback

Reija Helenius
Statistics Finland

The strategic objective of increasingly many organisations is customer orientation and development of activities through customer feedback. This paper presents how listening to customers and collection and processing of customer feedback can be systematised and included in the continuous improvement procedure. Systematic processing of customer feedback, emphasising the customer voice, can be supported in various ways in an organisation. Examples of this are year planner models for customer feedback processing and collection and introduction of indicators for customer effectiveness. Monitoring of customer feedback and experiences can be made by means of both conventional studies and inquiries as well as through new methods. Online brainstorms, validation of telephone calls and email services and studying of search engine

use introduce new viewpoints. By collecting together the feedback received from all different channels we gain an extensive image of customer feedback and needs. This equally includes monitoring of the use of web services and social media channels, media monitoring and data supplier feedback received. The paper presents practical tools and describes experiences of using various methods. As a case example, it is illustrated how the aims of listening to the customer in Statistics Finland's strategy have been advanced in practice. A key question in taking the strategy forward has been how the customer strategy is to be implemented in the whole organisation and its different levels. The paper produces an overall image of how customers are listened to, what ought to be measured and how to react to the feedback received. In addition, it is discussed how the collection and utilisation of customer data can be coordinated comprehensively in an organisation. The customer concept is considered to comprise stakeholders, data suppliers, data users and paying customers alike.

Keywords: customer feedback, customer strategy, continuos improvment

Title of abstract: Changing how we communicate quality and methods to our users

Sarah Tucker
Office for National Statistics, United Kingdom

How many of us can successfully build a flat pack piece of furniture without good instructions? Not many! Similarly, our data users need clear and concise quality and methods information to help them make better use of our data. At ONS, our products were previously designed to suit a paper format, heavily based on templates and containing lots of static information. With changing user needs and an emphasis on digital publications, we needed to develop new ways of communicating quality and methods to our users. We created a new strategy for communicating quality and methods based on meeting a wide range of user needs while being current, relevant and helpful to users. ONS analysts are encouraged to foster a sense of curiosity about their data so that they are able to share the most useful and meaningful quality and methods information that a) alerts users to potential limitations and b) helps users decide on suitable uses for the data. In this paper I will discuss how we used this approach to build quality and methods information into the structure of a new style of statistical bulletin which focuses on the story that the data is telling. This integral quality and methods information better enables users to make good decisions on the use of that data. I'll discuss the results of user testing that informed updates to our Quality Summaries and the creation of a new detailed Quality and Methodology Information report. I'll briefly discuss our Quality Information User Profiles which will help producers make decisions on what information to share. Finally, I will explain how our quality and methods products work together to create a layered approach to communicating the strengths and limitations of our data in a way that is accessible to all.

Keywords: Quality, users, communication, engagement

Title of abstract: PolicsLab: new data sources for a data informed policy making

Sara Laurita, Carmelofrancesco Origlia, Alfredo Fortunato, Mariagrazia Zottoli
Contesti srl, Italy

Statistical agencies are encouraged to explore the use of new data sources as input for official statistics, or combine them with existing ones, thus, weighing these opportunities against the need to maintain a high quality of statistical output[1][2]. Enrico Giovannini echoes this sentiment, stating: "Nowadays, NSIs are called to open the doors to new ways of constructing statistics in an era characterised by an infinity of new data sources"[3], such as big data, open data, and crowdsourced data (i.e. collaborative platforms). The advent of the data revolution has generated an overwhelming increase in the amount of data available, giving rise to new opportunities for policy-makers to use a data-informed approach in defining policies[4]. Official statistics data is based on international classifications and, although very useful, they are not sufficient to keep at the same pace with sudden changes affecting local territories, as well as external factors that may influence a given territorial policy. Despite the possibilities, new data sources are not necessarily a panacea for the actors involved in the policy cycle who have to tackle numerous methodological and ethical challenges[5][6]. This problem gave rise to the idea of offering users a concrete support tool to aid in the definition of policies by integrating various data sources that are useful for the adoption of an evidence-based approach. This paper presents the results of a case study in which we developed an „intelligent" platform for technological foresight called PolicsLab. The model is based on the valorisation of data available from traditional and non-traditional sources, as well as the construction of simple and intuitive features. The latter help to measure and predict territorial changes and transformations related to the innovation in Italian regions. PolicsLab is a solution for identifying appropriate analysis patterns, in order to help the user make an informed choice.

Keywords: PolicsLab, new data sources, data-informed, policy making support tool, innovation

Title of abstract: Ethnicity Facts and Figures – a marriage of good statistical practice and policy relevance

Richard Laux
UK Cabinet Office, United Kingdom

In October 2017 the UK Government launched an innovative website called Ethnicity Facts and Figures. Drawing together information held by government departments - survey and administrative data - it paints a compelling, and at times uncomfortable, picture of the different experiences of ethnic groups in areas as diverse as crime and justice, education, health, housing, and the labour market. This paper describes how we ensured that the website and overarching report comply with established statistical standards - for example:

Trustworthy: the decisions we have made – for example, about the range of data to be presented, the choice of time periods, the nature of the overarching analytical report, and quality assurance processes - reflected a clear commitment to transparency, accessibility and objectivity and were shaped by extensive user testing. Quality: in prioritising the data to be included in the Audit, the criteria we adopted included 'quality' and 'relevance'. We were guided by statistical experts across government about the quality of particular sources, and where possible we have drawn on official (including National) statistics. For each measure on the website we have included background sections covering a summary of noteworthy aspects of the data sources

and the associated methodology, and relevant web links. And we adopted quality assurance arrangements that made the most of the expertise of Departmental statisticians. Value: how we ensured that the statistics and commentary were accessible and relevant in helping the public to understand a very high profile issue - how people of different ethnicities are treated by public services. The paper also explores some of the challenges of maintaining statistical integrity in a project characterised by intense political interest and sensitivity.

Keywords: Ethnicity, policy relevance, Code compliance, user engagement

Title of abstract: Enhancing the awareness of official statistics in Egypt: The approach to increase their value

Ayman Hathoot
Central Agency for Public Mobilization and Statistics (CAPMAS), Egypt

Producing official statistics in The Central Agency for Public Mobilization and Statistics (CAPMAS) experiences an increasing demand from users. It should be met by independent, timely, impartial, trusted, high quality, and relevant data to support and enabling them to make a good decision. Enhancing the statistical awareness skills for data providers and users will achieve better use of it and increase their value. The paper demonstrates the approach for increasing in the official statistics value in CAPMAS through twofold. First, the applied approach to facilitate the use of an understood official statistics. It involves enhancing statistics learning at schools, raising statistical capability for targeted data provider groups such as employees in business and governmental sector, and enabling data users in accessing data and its related metadata. Second, the paper will expose to the Indicators used to monitor the progress of using official statistics which provide information on a specific aspect of the value of official statistics like the number of downloads, the number of citations by type of media, etc. (objective indicators) and indicators related user satisfaction (Subjective) to assess the value of statistics in terms of the user confidence and trust in official statistics, the usefulness and accessibility of official statistics.

Keywords: statistical capability - enhancing statistics learning – metadata

SESSION 13 – COLLABORATION WITH THE SCIENTIFIC COMMUNITY, COMMUNICATION WITH USERS

Title of abstract: Communicating official statistics and it´s quality in blogs – a rethorical analysis

Ann-Marie Karlsson, Jonas Hammarstrand
Swedish Board of Agriculture

The use of social media as an additional channel for communicating official statistics is increasing. Social media provides possibilities of highlighting and explaining statistics and builds the ethos of NSI:s for new groups of users. At the same time, texts in social media have a tendency to be oversimplified and overstated. I.e. social media places new demands on the NSI:s. In this article, rhetorical theory is used to analyse how the quality of official statistics is described in social media compared with how it is described in publications and databases. The case of how the Swedish Board of Agriculture, the organisation responsible for official agricultural statistics in Sweden, describes the quality of official statistics on consumption of food is chosen. A blog, a traditional publication and a database are compared. First, differences in the rhetorical situation, i.e. audiences, constraints and exigencies between the three ways of presenting statistics are highlighted. According to rhetorical theory ethos, pathos and logos are essential for building trust and the article therefore discusses how trust in the statistics might be affected by the use of social media as opposed to the more traditional channels for communicating official statistics. The findings show that the blog focuses on the consumption of meat, i.e. a type of food that is widely discussed in media. The blogposts are used by media and the public and are timely posted when meat consumption is debated in the society. The language in the blog is simplified compared with the language in the publication. The description of quality is mostly concerned with one aspect content. While the description of the quality was not as thorough in the blog as in the traditional publication, it was easier to comprehend and covered the aspects most essential to the content presented.

Keywords: rhetorical analyses, social media, blogs, ethos

Title of abstract: Collaboration with universities – a way to promote innovation and experimental culture at statistical offices

Faiz Alsuhail
Statistics Finland

Statistics Finland has had fruitful collaboration with the University of Helsinki and the University of Jyväskylä for over several decades. Thesis traineeships are an important part of this collaboration. Each year, a handful of students are recruited for a period of six months. The main task of the student is to write a thesis on a given topic. Statistics Finland provides the student with a research problem, data set, facilities and day-to--day guidance. The academic supervisor is from the student's university. This paper presents how a research problem is drafted in collaboration with the academia, how the students are recruited and supported during the traineeship as well as what the trainee program has resulted. We illustrate by concrete examples how the thesis traineeship can promote innovation and experimental culture in a statistical office. These

examples involve testing new tools and methods. We argue, that the thesis traineeship is a good way to involve the scientific community in the activities of a statistical authority. In particular, we show that the thesis traineeship is not only a mutually beneficial practice for the academia and statistical offices but also for the students.

Keywords: academia, traineeship

Title of abstract: How widening access to researchers has improved and can continue to improve microdata quality? From national experiences to European aspects

Kamel Gadouche
GENES-CASD, France

There are now a number of experiences at national level where opening access to the official microdata for researchers has been an important driver for improving microdata quality both for the metadata and for the data themselves which are two closely associated faces of data quality. Based on the French experience, the paper provides examples and discuss possible developments with increasing access to confidential microdata both at national and European level. Regarding metadata quality, one important aspect of the GSBPM (Generic Statistical Business Process Model) that claims as the DDI Lifecycle (it was not the case for DDI codebook 15 years ago) to managing microdata from end to end, is that it includes from the start to the dissemination phase researchers. Thus documentation doesn't remain the last thing to do after having produced microdata. Metadata have to be designed taking into account the reuse of microdata for scientific purposes which in turn increase the quality of metadata also for internal further use for the statisticians. In some cases, researchers can directly contribute to document microdata. This is particularly the case for administrative data. Since access to administrative data has been opened to researchers, many administrative data have been documented, sometimes with the help of researchers themselves (for example the tax data in France). Researchers' input on the data quality itself seems currently less systematically integrated though there are numerous examples of how feedback from researchers can be used. One obstacle is the lack of organized process to involve the researchers and collect these feedbacks. Opening more access to European microdata as well as allowing easier transnational access to official microdata for the researchers would certainly allow them to provide more input on the quality of data and metadata which are crucial for comparability, a core issue for public policies.

Keywords: metadata, quality, researchers, scientific, feedback, microdata, confidentiality

Title of abstract: Quality in statistics and relations to the press and media

Carsten Zangenberg, Steen Dahl Pedersen, Karin Blix
Statistics Denmark

As official statistical authorities, the statistical institutions have an important role in supporting a common fact-based perception of reality in society. Several of the traditional media also assume this role, and statistical institutions typically have formalised relations with the media. From a media perspective, however, the development in the past few years may be cause for serious concern. The internet has made everybody a publicist. Combined with the development in social media, it has changed journalism – and possibly also democracy. In the early days of the internet, the assumption was that it could contribute to strengthen our

democracy and the public debate. Instead, the amount of knowledge has become immense, and it is often impossible to decode who is behind the information and what characterises the quality. The internet has become the most monitored place in society, and it is controlled by giant global players, such as Google and Facebook. In addition, these players have won an increasing share of the advertising revenue that was previously an important source of income for traditional media. The challenge in this is emphasised by the fact that many media have not yet come up with a business model that can finance serious journalism and also work in a situation where online news take up ever more space. A question is whether the national statistical offices have adjusted their relations with – and interaction with the media so as to make opportunities as well as pitfalls in the use of the official statistics known. Another question is whether the media's use of the statistical institutions' work on the quality and the dissemination of it is utilised to an extent that bears comparison with the costs. In the paper, Statistics Denmark presents various initiatives taken based on these challenges and the effects achieved.

Keywords: Alternative facts, Quality facts, Media relations, Communicating qualityy statistics

Title of abstract: The dimensions of quality statistics in news. Strenghts and weaknesses

Alessandro Martinisi
Breda University of Applied Sciences, Netherlands

Statistics are often associated with the notion of quality and credibility, as they are seen as a way of neutrally measuring and operationalising abstract and deductive reasoning. Simple statistical vocabulary such as averages, percentages, modes, medians or common values and proportions are included in the greater part of news reporting. In journalism, statistics are treated both as a credible source of information as well as hard facts. It is assumed that by inserting quantitative information in a story, reporters can contextualise a news item into the wider framework of societal trends while at the same time use the statistics themselves to strengthen the argumentation upon a certain event. All in all, the broad understanding is that by incorporating statistics journalists can make their stories better quality. However, as I ask in this presentation, does statistics really improve the 'quality' of the news stories? Does quality statistics translate into quality journalism? Quality journalism has been long debated among media scholars especially in recent years. Little has been said however on how statistics contribute to the debate around the notion of quality journalism. The empirical research attempted to "measure" the concept of quality and then contextualise it into the journalistic practice. The presentation is drawn upon empirical data from content analysis and in-depth interviews among journalists in the UK and tries to reflect on the strengths and weaknesses about the articulation of quality statistics in printed news. In addition, the research underlines the dynamics of the data-analysis routine and the numeracy of those journalists who use an alternative voice through mathematical tools.

Keywords: Quality Statistics - Quality Journalism - Statistical literacy

SESSION 14 – VISUALISATION AND STORYTELLING

Title of abstract: Data Visualisation: Presenting compliance burden across GSS official statistics

Hannah Brunning
Office for National Statistics, United Kingdom

"Data! Data! Data! ... I can't make bricks without clay!"[1]. The importance of data permeates through history and as Doyle suggests, it forms the building blocks for which we base our conclusions. However, while the data affords us the opportunity to theorize, scrutinize, and make decisions, we are then faced with communicating our findings. It is here we realise the true significance of data visualisation. Data Visualisation is "the representation of data to facilitate understanding"[2]. The human brain is remarkable, but it has limitations. When it comes to data, especially that of a complex nature, it's much more digestible to communicate this data through graphs and charts. By appealing to our creative perceptions, data visualisation provides us with insight through aesthetically pleasing platforms. I will explore the role of a data viewer and what they need as they perceive and interpret data in various formats. This research is conducted surrounding a user review on compliance and burden across the UK GSS[3]. The multidisciplinary nature of this field offers the opportunity to look at both viewer and visualiser. There is a need to demonstrate trust and flexibility throughout the process of communicating data and this paper will explore the opportunities and challenges experienced. Trustworthiness is a recurring theme throughout; it's something which is necessary when presenting internal or external customers with statistics and it's a pursuit that should guide all decisions during analysis and design stages. Overall, I will demonstrate the transition from traditional data communication to data visualisation through compliance burden on all UK Government Departments who produce official statistics. I will also convey the importance of data visualisation, touch upon tools and technology as enablers and explore the analysis and communication of data through the perspective of both viewer and visualiser.

Keywords: Data visualisation, communication, theorize, scrutinize, and make decisions, insight through aesthetically pleasing platforms, multidisciplinary, analysis and design, technology

Title of abstract: Targeting a wider public – interactive storytelling with statistical data

Zsolt Czinkos
Hungarian Central Statistical Office

Reaching a wider public with statistical data is difficult. Smartphones and increasing mobile bandwidth are changing user expectations. Statistical visualizations should meet the standards of current user experience. Hungarian Central Statistical Office has started to create interactive storytelling infographics and data visualizations to highlight interesting facts, to explain terms, to show results – to improve statistical literacy. Creating customised story visualizations is challenging. It requires cooperation between people from different domains: software development, statistics, communication, management, visualization. It also needs software tools. Publishing to both mobile and desktop environment requires responsive design and cross-browser compatibility. Tools exist, but development is expensive. Is it worth it? Number of visitors should be measured, feedback should be received. This presentation offers insight into the development process of a published interactive storytelling visualization highlighting technical details.

Keywords: data visualization, storytelling

Title of abstract: Visual presentation of statistics on the Official Statistics Portal (OSP)

Laima Grižaitė, Jana Vanagė
Statistics Lithuania

Statistics Lithuania has several tools for statistical information visualizing. One of them is GIS tool. For a long time, Statistics Lithuania has only used geographic information systems (GIS) for drawing statistical maps for statistical publications. In 2017, the Official Statistics Portal, including the GIS component, was substantially updated. The GIS component was supplemented with additional functionality, enabling users to perform more complex analyses. On the Official Statistics Portal, two applications presenting statistical data in maps – Interactive Atlas and Detailed Statistics – were created using GIS technologies. The Interactive Atlas allows the data analysis in space and time, as well as comparison thereof. The Detailed Statistics application provides the indicators of the 2011 Population and Housing Census of the Republic of Lithuania at more detailed territorial level, i.e. by grids. The size of grids – 10, 5, 2.5 and 1 km2. The grids cover the whole territory of Lithuania; towns are divided into grids of 500 and 250 km2, the largest cities – 100 km2. The grids make it possible not to stick to administrative units and the collected data of that kind of detail in different periods always remain comparable. It is one of the first publicly available projects in Lithuania which allows receiving statistics of that kind of detail represented in maps. After the implementation of the Official Statistics Portal project and integration of GIS, Statistics Lithuania commanded widespread attention of the users and this encouraged to take on the development project.

Keywords: visualisation, GIS, interactive

Title of abstract: Bringing Data to life with Data visualization

Neal Chen, Jia Ying Goh, Wei jing Heng
Ministry of Manpower, Singapore

In today's data-rich environment where users are overloaded with vast amounts of information, data visualization is essential in capturing a viewer's attention and retaining it through storytelling. When data and visuals are weaved together into a story, it is able to resonate with the users on both an intellectual and emotional level. The human mind is able to grasp information better through data visualization tools instead of pouring over complex data in spreadsheets or reports. Although static visualization has long been used in storytelling, technological advancements have made the creation of interactive visualization easily accessible to the masses. Focusing on the users' experience as the centerpiece of our design, Ministry of Manpower (MOM) of Singapore has developed an Integrated Manpower Analytics System (iMAS) which offers a wide array of statistical products such as infographics, videographics and interactive dashboards. This paper examines how MOM integrates data visualization into our statistical production process to provide users with a unique and customized narrative experience.

Keywords: Data Visualization; Technology; Storytelling; Design

SESSION 15 – STATISTICAL EDUCATION AND LITERACY

Title of abstract: STAT WARS – May the force be with you! – Statistical competition for students

Mária Jónyer, Zsolt Kocsis-Nagy
Hungarian Central Statistical Office

The Hungarian Central Statistical Office (HCSO) has organised a national competition on statistics for secondary school students entitled STAT WARS for four years. Teams of 4 students and their teachers may enter for the competition, who start with completing a test of theoretical questions. In 2017, 154 teams applied from the country's various secondary schools, and the best twelve of them gained admission into the national final, which was held at the Headquarters of HCSO. The tasks of the STATWARS are built around a central topic every year, this year the focus was on the 150th anniversary of the establishment of the Hungarian statistical service. In the final of the competition, the 12 teams competed in 12 stations, and in each station, a theoretical and a practical task had to be completed in time, for which they could also prepare from the predefined recommended literature. During the tasks, among others, they calculated variance, completed a questionnaire about the subjective well-being of HCSO colleagues and could have a look into the significant statistical figures of recent centuries. In addition to the specific tasks, they could get acquainted with the history of the 120-year-old building and make a short film for the promotion of the 150 years of the office. The success of the statistical competition is indicated by the fact that some teams are re-entering for it year after year, and with the participation of 154 teams, we have reached more than 600 secondary school students who are interested in the statistics, as well as the event has been included in the programme 'TalentGate (Tehetségkapu)' of the Office of Education.

May the force be with you!

Keywords: education, statistical literacy, communication, target group, competition

Title of abstract: Promoting Official Statistics through Statistical Literacy Products

Maria J Vinuesa
National Statistics Institute of Spain (INE)

In this post-truth era we live in, the misuse of statistics and data is a fact. Its noxious effects are boosted if by chance they become viral in the social media. This is why now it is more important than ever to reinforce the efforts to increase the statistical literacy of our society. This imply effectively promote it as much as possible. This is exactly one of the aims of the DIGICOM project where INE-Spain and some members of the EU system are collaboratively putting together their best practises and successful promotional activities. An example of this is the European Statistical Competition.

Keywords: statistical literacy, promotion, best practices

Title of abstract: Experiencing Multidisciplinarity in a Master in Official Statistics

Jose Manuel Robles, Javier Montero - Complutense University, Madrid, Spain
Francisco Veira - National Statistics Institute of Spain (INE)

This paper focusses on quality controls within the Master in Official Statistics and Social and Economic Indicators at Complutense University of Madrid (UCM), particularly to assure multidisciplinarity. Being this master a joint project of 8 different Faculties, our global approach to Official Statistics counts with the support of the members of the National Commission of Official Statistics, leaded by the National Institute of Statistics (INE). The group of students it self represents an interesting heterogeneous team that indeed enrich the master. Besides a strict acceptance process, a personalized study program, regular student questionnaires and periodical staff meetings, Friday evenings are mainly devoted to complementary courses. They are organized upon demand of any member of the National Commission of Official Statistics. We were very soon asked to organize a course on communication of official statistics, developed in collaboration with a 9th UCM Faculty, the Faculty of Communication Sciences. Another course on gender in official statistics was organized in collaboration with the Ministry of Education (Government of Spain). Other courses were focused on advanced specific use of R language and SAS in official statistics. Moreover, in collaboration with INE, we have launched the series of Workshops in Official Statistics and Social and Economic Indicators (first two editions were held at the Faculty of Economics and the Faculty of Statistics). We have also launched a Workshop in Big Data and Social Networks and Social Analytics. Although the students coming from our EMOS Master have the preference to attend these extra courses and workshops, they are open to other students and professionals meanwhile the number to desired participants is not reached. We not only improve the master this way, but we attract students and we also help Society to be conscious of the key role that official statistics plays in their life.

Keywords: Education in Statistics, Science, Society, Statistical Institutions

Title of abstract: Improving labour market statistical literacy for organisations and individual users

Eugenia Goh
Manpower Research and Statistics Department, Ministry of Manpower, Singapore

Karl Pearson once said "Statistics is the grammer of science". Statistics is important and useful when utilised properly. As the National Statistical Agency for manpower statistics, the Manpower Research and Statistics Department (MRSD) of Singapore's Ministry of Manpower collects, analyses and disseminates essential statistical information on Singapore's labour market. Helping people understand and utilise manpower statistics is just as important as our primary statistical activities. To improve labour market statistical literacy, we have different approaches to reach out to the relevant stakeholders. This paper highlights the steps we adopt to improve labour market statistical literacy and reduce statistical misconceptions among relevant stakeholders, whom can be categorised into organisations (i.e. educational institutions, private corporations, the media outlets and the public sector) and individual users. To assess the level of labour market statistical literacy among the relevant stakeholders, we conduct informal polls to gauge users' awareness of the latest manpower statistical information and the concepts behind them. Our approach to improve manpower statistical literacy and reduce statistical misconceptions is tailored for each stakeholder. Our outreach

efforts include liasing with various organisations to bring statistical information to them. An example would be MRSD's collaborations with tertiary institutions. We conduct sessions with students at educational institutions to raise their awareness of the labour market situation and show them how such information could be used for their job search. For individual users, we publish regular papers and infographics that inform users of the latest labour market happenings and explain confusing concepts such as the concept of unemployment.

Keywords: labour market, manpower, statistical literacy

Title of abstract: Statistical literacy and education in the State Statistical Office of Macedonia

Mira Todorova
State Statistical Office, Macedonia

In modern knowledge-driven society, the knowing how to use statistical information is a necessary skill to citizens. The State Statistical Office of Macedonia recognizes the importance of the proper use of statistical data. According to its Strategy, State Statistical Office will work on improving statistical literacy and the specific sub-program: Improving statistical literacy is defined. In the paper will be described the actions taken by the State Statistical Office to promote the statistical culture in the State. Besides, it will be given the description of post- graduated studies on „Statistical methods for Business and Economy", held at Faculty of Economics in Skopje.

Keywords: Post-graduated studies, statistical literacy, SSO Strategy

SESSION 16 – MEASURING AND COMMUNICATING QUALITY

Title of abstract: A Future Focused Approach to National Statistical Quality Reviews in the UK

Gentiana Roarson
Office for National Statistics, United Kingdom

The production of official and national statistics requires a programme of continuous improvement and periodic reviews to ensure that the outputs provided remain helpful, trustworthy, and represent value for money. The Office for National Statistics in the UK is developing and implementing a new approach to National Statistics Quality Reviews (NSQRs) aiming to monitor and improve the quality of statistical outputs in–line with data revolution and changing user demands. The new NSQRs are future focused, cover themes of national importance and are conducted by multi-disciplinary teams across the Government Statistical Service (GSS), academia and the user community. Each NSQR will incorporate two phases, an initial phase to cover activities of the review itself and the formulation of recommendations and, a subsequent phase to follow up arrangements for recommended actions and post implementation evaluation. The fist NSQR, in this new format, is a review of the privacy and data confidentiality methods aiming to highlight areas that require development and investment to help producers of official and national statistics to prepare for the future. This review is carried in the context of new legislation being introduced in the UK to allow statistical producers across GSS to use non-survey data for the purpose of producing more detailed, frequent and flexible statistical outputs. The use of richer data sources poses new challenges in ensuring that privacy of individuals and confidentiality of the data is protected at all times. This presentation will discuss both the challenges and the benefits of implementing the new approach toquality reviews.

Keywords: national statistics quality reviews, data confidentiality methods, privacy methods, non survey data, continuous improvement

Title of abstract: An empirical study of the effect of high, low or moderate nonresponse rate on the quality of survey estimates

Øyvin Kleven, Ib Thomsen - Statistics Norway
Li-Chun Zhang - University of Southampton, United Kingdom

Response rates have decreased steadily in household surveys across many countries, despite the increasing effort and resource being spent to deal with the problem. The associated nonresponse errors have received considerable attention in the past decades, as they can cause bias and are critical to the accuracy of survey based statistics. Some authors have published papers claiming that estimates of sufficient quality can be produced based on response rates as low as 5 percent (e.g. Hellevik 2015), while practitioners and users of survey data often argue that "high" response rates are necessary for god quality. In this empirical study we focus on the Norwegian Election Survey data over nearly 50 years, combined with linked administrative sources, which provide additional data of Electoral Turnout and population demographic characteristics. The nonresponse rate has increased steadily from 10 percent in 1969 to 45 percent in 2017. But how has nonresponse bias evolved over the same period? Can another indicator better capture the nonresponse bias, such as the R-indicator that has received much attention in the recent years? Is it possible to devise other simple

but more useful bias indicators? We take a closer look at the absolute deviation of the response rates as an alternative to the squared deviation used in the R-indicator. How would the nonresponse bias evolve, based on extrapolation of the trend of nonresponse rate and other factors that determine the nonresponse bias? What would be the likely nonresponse bias if the nonresponse rate either becomes very high or very low? Is there an "acceptable' range of nonresponse rate in practical surveys? These are some of the questions we investigate based on the historical material at our disposal.

Keywords: Survey nonresponse, Bias, Indicator, Historical data

Title of abstract: Measuring, reporting and communicating quality of National Accounts statistics (ESA 2010) in an integrated way with data production

Ani Todorova, Christos Liouris
Eurostat

The Regulation (EU) No 549/2013 sets up the European System of Accounts 2010 (ESA 2010). Article 4 of this Regulation requires Member States to provide to Eurostat quality reports which would allow the Commission to assess the quality of data received under ESA 2010 transmissions. The Commission Implementing Regulation (EU) 2016/2304 lays down the modalities, structure, periodicity and assessment indicators of the quality reports. Under the above mentioned regulations Eurostat in collaboration with the Member States has set up quality reports for National Accounts. To set up a business process, specific challenges needed to be addressed. Statistical production processes routinely generate large amounts of information of potential interest for quality reporting. Quality reporting has however traditionally been carried out as a parallel and separate activity from data production, and has therefore often not made full use of the wealth of information created by production processes or stored in production databases. The paper, firstly, introduces the framework in which the annual quality reports are produced. In the second part the quality measures are presented. The third part shows Eurostat's efforts to automatically integrate information coming from data production into the quality reporting process. The fourth part focuses on the implementation and the communication of the country reports and the Commission's report to the European Parliament on the quality of National Accounts. Though the quality reports cover the full transmission programme under the ESA 2010, examples in this paper will concentrate more on the domain of main aggregates and supply, use, input-output tables. Finally, the paper highlights the challenges encountered and discusses future avenues of work to achieve more efficient ways of quality reporting.

Keywords: measuring quality, quality reporting and communication, national accounts, ESA 2010, SDMX, SIMS/ESQRS

Title of abstract: Quality Management for Official Statistics: Some Lessons Learned after Seven Years of ASPIRE

Paul Biemer - RTI International and University of North Carolina at Chapel Hill, United States, Heather Bergdahl, Joakim Malmdin- Statistics Sweden, Dennis Trewin -Melbourne, Australia, Dan Kaspyrzk - NORC, Washington DC

ASPIRE (A System for Product Improvement, Review and Evaluation) provides a general framework for improving the quality of statistical programs for National Statistical Offices (NSOs) and other organizations that provide a continual flow of statistical products to users and stakeholders. ASPIRE was first implemented at Statistics Sweden in 2011 in response to a mandate from the Swedish Ministry of Finance to develop a system of quality indicators for tracking developments and changes in product quality and for achieving continual improvements in data quality across a diverse set of key statistical products. In this presentation, we provide an overview of ASPIRE, demonstrate its application to several statistical products (such as, for example, the Swedish Labour Force Survey, Consumer Price Index, and Structural Business Statistics) and summarize some of the key lessons learned after having applied the system annually for the past seven years. We will present some results from an organization-wide evaluation of ASPIRE (at Statistics Sweden) which has led to some important modifications of the ASPIRE process going forward. The presentation will also discuss the implications of these changes to ASPIRE for monitoring and evaluating product quality at Statistics Sweden as well as in statistical organizations world-wide.

Keywords: quality assurance, statistical standards, total survey error, data accuracy, quality control

Title of abstract: Automatically Generated Quality Control Tables and Quality Improvement Programs

Justin Nguyen, Carma Hogue
United States Bureau of the Census

The Economic Directorate of the U.S. Census Bureau collects various economic data with a requirement to accurately capture, process, and analyze the data to guarantee our processes identify and correct quality issues. Effective quality control systems are the foundation for successful data collection along with well--defined program requirements that assure compliance with Office of Management and Budget and Census Bureau quality standards. To meet these objectives, our primary goal is to build an automated quality control and quality assurance system that will identify and implement analytical methodologies that reduce the introduction of error into analytical data. For data collection and data evaluation, we intend to build a system that ensures that all surveys conducted by the Economic Directorate produce the quality results expected for the intended use of the data. In this paper, we discuss requirements for a set of automatically generated tables and applications that should be used to monitor various processes from the survey planning stage through product dissemination. In the remainder of the paper, we discuss the automated quality assurance checks that should be made in each survey phase to ensure that decisions will be supported by data of adequate quality and usability for their intended purpose, and further ensure that such data are authentic, appropriately documented, technically defensible, and statistically sound. Quality issues and resolutions at each stage of the survey will be discussed. Likewise, resource challenges and possible solutions will also be addressed.

Keywords: U.S. Census Bureau, quality control, time interval, statistical period cycle

SESSION 17 – QUALITY CHALLENGES FOR CENSUSES

Title of abstract: Record linkage methods for Admin Data: Portuguese Census Transformation Program

Sandra Lagarto, Anabela Delgado - Statistics Portugal
Rui Silva, Lufialuiso Velho, Mário Silva, Pável Calado - INESC, Lisbon, Portugal

From 2014 to 2016 Statistics Portugal (SP) has been studying the usability of available administrative data for census purposes. Unlike other countries that have already made the transition to register-based or combined census models, in Portugal there is neither a central population register nor a unique personal identification number. SP created a methodological framework to build a Statistical Population Dataset (SPD), integrating linked registers in a "potential" resident database, then applying a 'signs of life methodology' to estimate residents [1]. For 2015, the population estimated from the SPD was 10,434,161 persons, with a deviation of 0.9% to the official resident population estimates for the same year. Results are promising at a national level, but still, there are multiple hurdles to the creation of this dataset that don´t particularly benefit the results at a detailed geographical level: records have inconsistencies and errors due to manually inserted data, and the National Data Protection Authority imposed anonymization criteria on the datasets, restricting access to the full name and address of the persons in registers. Exact comparison methods performed by SP left out many potential matches (roughly more than 5% for most sources). With the goal of identifying the highest number of linked pairs of records, a cooperation a between SP and the University proposed an alternative linkage model [2], based in a logistic regression, which added thousands of new pairs of linked registers to those found by SP. The precision of the method is about 99% on a large set of linked records used as gold standard.

[1] Statistics Portugal. 2016. Metodologia atualização Base População Residente – Construção BPR 2015. 2021 Census Unit.

[2] Silva, R., Sampaio, L., Calado, P., Silva, M., Delgado, A. 'Matching administrative data for census purposes'. In Proceedings of the Data Science, Statistics and Visualization Conference, Lisbon, Portugal, July 2017.

Keywords: Census, administrative data, Record Linkage, Machine learning applications

Title of abstract: Assessing quality in a register-based census

Marina Pérez Julián, Antonio Argüeso
National Statistics Institute of Spain (INE)

Spain is going to conduct for the first time a Census which will be largely based on registers. It will be built around the Population Register and will integrate administrative data from many sources. In order to help users to understand the production process and to better interpret the results, for each variable, another categorical variable measuring the quality of the information will be disseminated. These variables provide a powerful way for assessing quality in the census data across columns (variables) and rows (people and households). As the process for elaborating the census 2021 is ongoing, the intermediate results of this analysis is also helpful for focusing our efforts. In this paper we present some preliminary results based on the first general test of integration of sources we have done so far, the so-called pre-census file 2016.

Keywords: Quality, register-based census, integration of data sources

Title of abstract: Quality as a priority of 2021 Population census?

Elzbieta Gołata
Poznan University of Economics and Business, Poland

In the last 2010 round of censuses, we have witnessed historic transformations of unprecedented scope. The main motivation for changes in census methods was cost reduction. The arguments also indicated willingness to improve quality of estimates, timeliness and form of data dissemination due to introduction of modern technologies (UN, 2013). The rolling census methodology was first applied in France. In the United States, the American Community Survey replaced traditional long form. Alternative census methods using administrative registers increased interest. This approach stimulated development and application of modern methods of imputation, statistical data integration, calibration and small area estimation. The process of changes is not finished and it is difficult to predict its consequences. The discussion is lively, and some findings suggest ‚the end of censuses' (Baffour, King, & Valente, 2013; Coleman, 2013; Poulain & Herm, 2013; Valente, 2010). However, survey conducted for 1985 - 2014 indicated that the number of countries carrying out census increased, and methodology transformation concerns Europe (Kukutai, Thompson, & McMillan, 2015). The aim of this study is to present changes in methods of conducting censuses from the perspective of quality assessment. It attempts to show the scale and diversity of proposed solutions indicating their possibilities and limitations. Assessing quality, references were made to significance of censuses, their role for state administration and international comparisons (Baffour et al., 2013, p. 409). In accordance with Brackstone's definition (1999, 2003), all aspects of statistical survey reflecting usefulness to users were analysed. Similar, utility concept of quality was developed for official statistics and international organizations (Eurostat, UN, IMF, OECD), though components and classifications differ (Vries, 2002). The study took into account following six dimensions: relevance, accuracy, timeliness and punctuality, accessibility and transparency, comparability and coherence (Eurostat, 2003). The analysis concerns Polish experiences and perspectives, but includes references to other countries.

Keywords: census, quality, administration registers

Title of abstract: Selected Innovation Elements in Census 2021

Ludmila Ivančíková, Helena Glaser-Opitzová
Statistical Office of the Slovak Republic

The next population and housing censuses (the Census) taking place in Slovakia in 2021 is, in a way, unique. The change of concept is its essential feature. It is a transition from traditional population and housing censuses to combined ones based on the integration of data from registers and administrative data sources and data obtained from population. The implementation of the new concept covers several elements of innovation, including full electronic data collection, systematic evaluation of quality of administrative data sources used for statistical purposes, coordination of the Census preparation via the National Action Plan and data integration. The paper is focused on the presentation of those innovations that are specific to that change and are substantial to the preparation of the Census, which is labelled, in conditions of Slovakia, by the term "integrated Census".

Keywords: census, innovations, data integration

SESSION 18 – LABOR MARKET STATISTICS – NATIONAL EXPERIENCES

Title of abstract: The „EUROPA 2020" LFS indicators on regional level in Poland. How to improve the indicators precision without the increase of the LFS sample?

Hanna Strzelecka, Agnieszka Zgierska
Statistics Poland

In case when there is a need to evaluate some phenomenon by selected characteristics or by territorial division, the often encountered problem is finding an appropriate data source which delivers reliable results. Two of the five "Europe 2020" headline targets are monitored with EU Labour Force Survey (LFS) indicators. Moreover, several other indicators covering different EU policy domains are based on the EU-LFS results. Every Member State has to disseminate comparable and reliable data at regional level for regional policies monitoring needs. How to do this when the data are based on a sample survey, like the LFS, and response rate has been decreasing over time? The objective of the research work carried out in Polish Central Statistical Office (CSO) was division of main LFS indicators (not only "Europe 2020") into three groups:

1) ready for publication at regional level;

2) maybe possible - the precision of which may be expected to be improved to the acceptable level after the introduction of the significant methodological changes into the survey;

3) not possible at all for dissemination due to a low precision.

For indicators from second groupseveral simulation tests were done.

The two most important effects were achieved within the framework of the research work, which significantly have developed possibilities of the public statistics as regards the publication of the LFS regional data:

1) Specification of the base of indicators (from the list being the object of the research work) together with the codes of precision/quality embracing the historical data;

2) Elaboration of the variants of changes in the survey methodology, allowing to extend significantly possibilities for meeting the information requirements included in the scope of the research work.

Keywords: labour market statistics, Labour Force Survey, sample, precision

Title of abstract: Quality Improvements using indicators, business process improvement and change management

Fiona O'Riordan
Central Statistics Office, Ireland

Business Process Improvement in LFS. The CSO has moved recently from the Quarterly National Household Survey (QNHS) to the Labour Force Survey (LFS). Significant changes have been made during this transition including the use of CATi for later waves, a new IT System to monitor and measure metadata and paradata and a separation of the LFS from the General Household Survey (GHS). The GHS is now the instrument to collect the once off surveys like ICT, Crime and Victimisation and other European and National Surveys that are done on an ad-hoc basis. This was a large project for CSO and has had many challenges over the last

number of years. The LFS will go live on Q1 2018 and because of the challenges it has become part of the work of the Quality Assurance to review the process from questionnaire design to final publication. This paper will detail how we reviewed the processes of the many teams involved and how the quality division in conjunction with the relevant teams changed some of the older QNHS practices to fit the new LFS data lifecycle. The paper will detail how we assimilated the data, how the relevant teams involved produced some of the solutions themselves and how when we put all the existing processes together the teams had immediate suggestions for improvement. Quality indicators and risks were identified at a number of stages and this were built into the process control. Change management was key to this process improvement and the Quality team managed the change very well. The improvements were detailed and celebrated and this certainly gave momentum to continue the improvement. This paper will detail the various stages of change and subsequent improvement.

Keywords: Business Process Improvement, Quality, Indicators, Change Management

Title of abstract: Analysis on nonresponse bias for the Swedish Labour Force Surveys (LFS)

Martin Axelson, Vanja Hultkrantz, Fredrik Olsson, Pär Sandberg, Frida Videll
Statistics Sweden

All statistics have some degree of uncertainty that have an effect on the accuracy. The focus in this paper will be on object nonresponse and its effect on the nonresponse bias in the Swedish Labour Force Surveys (LFS). An analysis has been conducted by approximating LFS variables with variables from different registers. The variables being analyzed are; employed, unemployed, not in the labour force, employees, students, group after income and not in employment, education or training (NEET). Two different estimates are produced for each register variable by using the General Regression Estimator currently in use in the Swedish LFS, one by using the respondents and the other by using the sample. These estimates are used in order to estimate bias, relative bias and confidence intervals for the biases. All estimates are produced for the total population aged 16-74 and for relevant subgroups. Results for age group 16-74 in December 2015, age group 16-24 in year 2014 for NEET, show that the relative bias is significantly different from zero for almost all of the analyzed register variables, the exception is unemployed. Looking at the relative bias for employed, unemployed and not in the labour force over time, year 2011-2015, one can see that it has been relatively stable over the time period. Therefore, there is no clear indication of an increasing nonresponse bias as the nonresponse increases. In order to analyze the nonresponse in estimates of change, estimates of change comparing corresponding months from one year to the next have been computed for employed and unemployed for year 2013-2015. For unemployed, the estimates based on the respondents are similar to the estimates based on the sample. The same pattern is in general also seen for employed with exception for subgroups after education.

Keywords: Swedish Labour Force Surveys, Nonrespons bias, General Regression Estimator, Register variables

Title of abstract: Coherence between surveys and based-register data in labour market statistics

Sara Carrascosa, Laura Grande
National Statistics Institute of Spain (INE)

The study of the labour market has always had a special interest in society, both economically and socially. In Spain, there are different sources of information that allow us to approach and measure this reality. Statistics based on surveys are available, such as the Labour Force Survey (LFS) - carried out by the National Statistics Institute of Spain (INE) -; and others using administrative data, such as those based on the social security affiliation register and the unemployed persons registered in public employment services, provided by the Ministry of Employment and Social Security. Therefore it is interesting to study the coherence among these sources of information as an aspect of improving the quality of statistics. Under the framework of the High Council on Statistics the working group of Short-Term Labour Market Statistics draws up periodically a report with the aim of analyzing and comparing the data provided by the LFS and those provided by the Ministry of Employment, using for that purpose, a harmonized methodology. In addition, a series of short-term analyses are also carried out to study the coherence between them from the point of view of the final published data. This project allows, on the one hand, to compare, study and review methodological differences, and on the other to analyze the information that users receive from different sources of information.

Keywords: coherence, administrative data, labour market statistics

Title of abstract: On the exciting path to add a mode – redesigning the Austrian Microcensus Housing Survey into a multi-mode questionnaire

Vlasta Zucha, Barbara Leitner
Statistics Austria

The Microcensus Housing Survey is currently being redesigned to meet the growing demand of a multi-mode capable questionnaire. Besides computer assisted personal interviewing and computer assisted telephone interviewing, respondents will have the option to fill out the questions online. Although three different interviewing methods will be offered, the survey will be based on only one common questionnaire which needs to be highly user-friendly for all modes. Planning and redesigning the new survey instrument, specific challenges have to be taken into account: (1) the Housing Survey is jointly conducted with the Labour Force Survey within the framework of the Microcensus, (2) the output is linked to other statistical products, e.g. the Consumer Price Index, (3) the data is collected continuously throughout the year, (4) the selected households are interviewed quarterly and (5) the households have to answer questions on housing and housing costs, whereby precise information is indispensable. To implement a new mode means to develop new questions and to change the data structure. At different stages of the project, specific quality assurance measures are applied to ensure an optimal transformation from the old to the new survey instrument and to minimize breaks in time series. First of all, respondent debriefings as well as interviewer debriefings were carried out. Different data sets on housing were analysed. Considering all the analysis, a draft questionnaire was elaborated and tested by using qualitative methods. This questionnaire is now being redesigned. In mid-2018 a quantitative test of the final multi-mode questionnaire is planned.

Keywords: multi-mode data collection, web-survey, Microcensus, housing

SESSION 19 – QUALITY STANDARDS FOR STATISTICAL DOMAINS

Title of abstract: Quality of ICT skills indicators

Maja Rynko - Warsaw School of Economics, Warsaw, Poland
Marta Palczyńska - Educational Research Institute, Warsaw, Poland

Basic ICT skills are a prerequisite for social inclusion in the information age. Level of population ICT skills as well as computer and Internet usage are of interest to the European Commission and the national governments. This is reflected in numerous strategic documents (e.g. Digital agenda for Europe, A new skills agenda for Europe, Long-term National Development Strategy for Poland). The EU survey on ICT usage provides a proxy on the level of ICT skills among individuals in Europe. However, it is based on individual's self-assessment which may be prone to misjudgment (e.g. Dunning et al. 2004). Although potentially more accurate, direct assessment of skills poses a big challenge to survey design, fieldwork and budget planning. Using data from the Polish follow-up study on the Programme for the International Assessment of Adult Competencies (postPIAAC) we compare self-assessment of basic ICT skills to their direct assessments. The postPIAAC was conducted between Oct. 2014 and Feb. 2015 on over 5 thous. respondents aged 18-69. The basic questionnaire included Eurostat questions on computer and Internet usage. Respondents were also asked to conduct several tasks directly comparable to the Eurostat questions: coping files to a folder, using copy/cut/paste tools for text editing and using basic spreadsheet formulas. Among those who declared to have performed these tasks before, many did not complete the tasks in the direct assessment (13%, 8% and 65% respectively). The preliminary multivariate analysis suggests that young and higher educated people are more likely to overrate their skills in the self-assessment. The discrepancy between the actual and the declared level of ICT skills poses a question of reliability of ICT indicators based on self-assessment.

Keywords: ICT skills, quality of indicators, self-assessment, direct measurement of skills

Title of abstract: The method of harmonised Labour Market Areas in Europe

Valeriya Angelova-Tosheva, Dario Buono, Teodora Brandmueller
Eurostat

The Labour Market Area (LMA) is a well-established and discussed concept in regional geography and statistics. The concept of LMAs has existed for almost 30 years, based on different definitions and known under various names (Labour Market Regions, Employment Zones, Commuting Zones, Travel-To-Work-Areas, Daily Urban Systems, Working Catchment Areas etc.). The need for functional geographies in the statistical practice is indisputable. Administrative boundaries often break up single LMAs. Commuting across NUTS and country boundaries can lead to significant differences between total employment (job-place-based) in a region and resident working population (domestic employment) in the same region. Indicators such as GDP per inhabitant will be affected in regions with asymmetric commuting patterns. Luxembourg, Inner London and Brussels are only a few examples of territories where employment and GDP data are distorted when presented as divided by inhabitants. For the last ten years, Eurostat has intensively worked on the EU-wide harmonisation of the LMAs concept starting with a study together with the research community to

investigate the value assed, feasibility and best practices in the EU. A Task Force continued the work to make an official proposal of the Commissions position on harmonised LMAs. The approach for delineation of LMAs proposed to the Member States is a simple, transparent, reproducible, consistent, and policy independent bottom-up method that needs only commuting flows as input. As a following-up action, in the frames of a grant programme several countries tested the IT tool and proved the feasibility of implementing the harmonised method. This article aims to present and discuss the method in details, as in the future it will be the basis of creating LMAs as an operational service in a joint ESS initiative.

Keywords: Functional geographies, LMAs, harmonisation, R, self-containment, small vs. large LMAs

Title of abstract: Improving the quality of Business Statistics through Profiling

Ioannis Xirouchakis - Eurostat
Vincent Hecquet - INSEE, Paris, France

Economic globalisation and the way multinational groups organise themselves have led to increasingly complex organisations and to a growing gap between their legal and economic structures. Equating the statistical unit to the legal unit has resulted to two biases: to the impression that new enterprises have been created and to double accounting for non-additive variables (such as turnover). Such biases affect the quality of economic statistics. Confronted with this common problem, several European national statistical institutes decided to go beyond the legal definition of the legal unit and to implement enterprises in the economic meaning. This distinction is commonly done by „profiling". National profiling increases the relevance and quality of business statistics. First, profiling of the most significant groups relies on establishing communication between dedicated teams within statistical institutes and group representatives. Besides, it aims at suppressing „double-counting" for non-additive variables and at delineating the economic activity of groups on the national territory. Large groups are particularly internationalized, having affiliates in different countries. Therefore, cooperation between statistical institutes is required to understand their business and structures. Since 2013 Eurostat has been supporting projects with Member States to test European collaborative profiling and to agree on a common methodology based on test results. In 2018 more than 300 multinational groups have been collaboratively profiled. The main results in terms of quality improvements include (I) consistency of the attributes of multinational groups across national borders; (II) better understanding of the groups' activities, contributing to their consistent view and treatment; (III) more accurate evaluation of their footprint on the national economies and consequently (IV) improved quality of national statistical business registers and of the EuroGroups Register containing the biggest multinational groups present in Europe. Finally, the paper investigates examples for improving data and (subsequently) statistics quality using national and European Profiling.

Keywords: business statistics, profiling, multinational groups, enterprises

Title of abstract: Profiling: a new way to increase the quality of statistics on research and development

Thomas Balcone, Camille Schweitzer
Ministry of Higer Education, Research and Innovation, France

Currently, statistics on Research and Development (R&D) carried out in the business sector are computed in France on the sole basis of legal units: firstly, a survey is addressed to them to collect the data and then, statistics on R&D are disseminated at legal unit level. Considering the increasing importance of the enterprise group in the French economy, it seems difficult today to go on using only the legal units to calculate business statistics. Indeed, assimilitating the legal unit to the enterprise is not relevant anymore for group's affiliates and subsidiaries. Taking into account the European definition of an enterprise will help to disseminate more consistent and relevant R&D statistics on the business sector. The French business statistic register established by the French national statistical institute (INSEE), called SIRUS, contains notably all the legal units, all the enterprises and all the links between them. The main contribution of this register is to make possible the calculation and dissemination of statistics at an other level than the legal unit one: the enterprise level. This article first describes why the data should go on being collected at legal unit level and not at enterprise one. Indeed, it seems that such a change in the data collection can be dangerous because it could result in a substantial increase of the response burden. Then, this article presents the process based on SIRUS that leads to the computation of key indicators on R&D at enterprise level. To conclude, it compares these key indicators with the ones calculated at the legal unit level to show the impact of moving to the enterprise level on French R&D statistics.

Keywords: Business register, research and development statistics, statistical unit

Title of abstract: European structural farm statistics – new quality rating system

Denisa Florescu
Eurostat

Eurostat, together with statistical bodies in the ESS, have adopted a quality rating system that guides how structural farm statistics derived from farm structure surveys are disseminated. The system does this by showing when the estimates are sufficiently reliable to be published, with or without warning. It is based on:

• coefficients of variations for totals and means of continuous variables;
• standard errors for proportions and counts.

The paper will also present the work carried out to harmonise methods and application of variance estimation methods within the ESS. To consistently apply the new quality rating system, Eurostat and the national statistical bodies must compute roughly the same variance estimates. Future structural farm statistics will come from the data collected from 'Integrated Farm Statistics', based on a modular approach. This will lead to more complex national sampling designs. Also, the paper will outline ongoing developments to integrate additional sampling design information specific to national multi-stage sampling in the estimation of variance. It will aim to accurately compute variance in Eurostat. The paper will also introduce new quality reporting based on the European Standard Quality Reporting System (ESQRS) template. This will greatly help to assess all quality dimensions, improving the quality of EU data and metadata. Farm structure surveys are the main source of information on the current state of and trends in agriculture required to monitor the common agricultural policy and other EU policies. High-quality data is essential for decision-makers.

Keywords: farm structure surveys, quality, quality rating system, variance estimation, quality reporting

SESSION 20 – QUALITY IN THE FUTURE – THINKING FORWARD

Title of abstract: Quality of official statistics in micro and small economies in globalized world

Józef Oleński
Lazarski University, Warsaw, Poland

The majority of economies and countries in the world of today (over 100) are micro and small economies (MSE). In globalized world the prerequisite of social and economic policy of governments of those economies and countries is the availability of complex, reliable, relevant and pertinent information adjusted to the specificity of each particular micro and small economy and relevant integrated data on their political, social, economic and ecological environment. The production and dissemination of such information for governments, social organizations and businesses is the duty of national systems of official statistics of the MSE. Micro and small economies are insisted by international organizations to submit statistical data following exactly the international standards that are adjusted to the specificity and the needs of large countries. However, those standards, methods, categories and data are often difficult for compiling, hardly interpretable or even useless for MSE. The objective of this paper is the analysis of social and economic attributes specific of micro and small economies that have the impact on information sources, methods and needs of national users of statistics. The typology of micro and small economies is proposed. The consequences of the specificities of different types of the MSE for the duties, functionalities and strategies of development of official statistics and problems of statistical capacity building for the MSE are discussed. In the conclusion it is formulated the need of creating the international, joint research institute providing scientific and educational assistance to national statistics and educating higher level statisticians (MOS - Masters in Official Statistics) for all micro and small economies.

Keywords: official statistics, micro and small economies

Title of abstract: The future role of international organisations: enhancement of relevance and quality of official statistics

Maria Joao Santos, Cristina Pereira de Sá
Eurostat

The authors analyse in the paper the future of official statistics, focusing in particular on how the international organisations through changes in governance or standard setting could contribute to increasing its relevance and quality. Official statistics are faced nowadays with a multitude of challenges (e.g. shrinking resources, improvement of timeliness, need to give responses of greater relevance to the specific requests of users, use of new data sources) and have to provide an effective answer through high quality official statistics in many domains at global level (e.g. SDGs, globalisation). Under these rapidly changing circumstances, it will however be essential to enable Official Statistics to play their important societal role through appropriate adaptation of the rules, the principles and resources, which frame their working conditions. Enhanced cooperation and communication between the main stakeholders at global level is essential for improving

the credibility and relevance of official statistics. But in order to be able to face the current challenges the global statistical system needs to go a step further. The authors will discuss the feasibility of the application of the European Statistical System model to the global statistical system by considering the application of the ESS standard setting model and the transfer of aspects of the ESS governance and cooperation structure to the global statistical system. Is this possible without an underlying legal framework as it exists in Europe? Is there room for strengthening the agreements at global level and establish new cooperation mechanisms? In this context the authors will argue for a change in international statistical governance and address issues such as the new role of international organisations.

Keywords: Future official statistics, cooperation with international organisations, relevance, ESS

Title of abstract: Conceptualising quality for big data

Giovanna Brancato, Loredana Di Consiglio
Italian National Institute of Statistics (ISTAT)

Whereas the importance of big data for official statistics is widely recognised, the quality of the big data is a concern and the statistical community has soon started to reflect upon a framework to assess their quality. For instance, the United Nations Economic Commission for Europe (UNECE) Big Data Quality Task Team (UNECE 2014) extended an administrative data quality framework; the AAPOR total error framework for big data (Kreuter, 2015) focused on an extension of the total survey error examining sources of errors specific for big data. Reis et al. (2016) have analysed a few alternative approaches on real case applications, concluding for the need to complement the different approaches and to structure the links between input and throughput quality and output quality. In the Essnet Quality of Multisource Statistics, a framework relating the output quality with the sources of errors is being proposed (Brancato, 2017). The proposal takes into account the main quality models for administrative data. This paper aims at investigating the applicability of the Brancato approach for big data and proposes the necessary extensions taking into account of the quality frameworks for big data. Finally, very preliminary thoughts on the relevant factors when producing experimental statistics are given.

References

- Brancato, G. (2017) Unpublished technical report WP1. Guidelines on the quality of multisource statistics, Essnet on Quality of Multisource
- Kreuter, F., Berg, M., Biemer, P., Decker, P., Lampe, C., Lane, J., ... & Usher, A. (2015). AAPOR Report on Big Data (No. 4eb9b798fd5b42a8b53a9249c7661dd8). Mathematica Policy Research.
- Reis F., Di Consiglio L., Kovachev B., Wirthmann A., Skaliotis M. (2016) Comparative assessment of three quality frameworks for statistics derived from big data: the cases of Wikipedia page views and Automatic Identification Systems, Q2016
- UNECE (2014) A Suggested Framework for the Quality of Big Data

Keywords: quality hyperdimensions, sources of errors, statistical process

Title of abstract: A knowledge-based approach to the statistical production process

Stefano Menghinello, Claudio Ceccarelli, Massimo Fedeli
Italian National Institute of Statistics (ISTAT)

The debate on data quality in official statistics has recently shifted its focus from an output oriented to a production process perspective. This paper explores the link between the design and management of statistical production processes and data quality measurement and evaluation in the light of contributions from the knowledge economy and knowledge management literatures. The classification of statistical production either as a knowledge transfer or as a knowledge intensive process is crucial. In the first case, the knowledge embedded in data collected from respondent units is shifted to final users with no alteration in data processing, while in the second case the knowledge is transformed to generate more consistent, accurate and business relevant data. The knowledge economy literature tends to acknowledge that any production process is knowledge based, since the knowledge embedded in raw inputs (being goods or services) is either explicitly (codified knowledge) or implicitly (tacit knowledge) transformed into the final output by the human capital and technologies engaged in the process. The aim of this paper is twofold. Firstly, it highlights the key features of statistical production as a knowledge intensive process, also providing some concrete examples in the broad domain of globalisation, including recent development on profiling, establishment of large case Units and EU level early warning system. Secondly, it exploits the consequences of this approach on data quality with respect to different features of business survey management: questionnaire design, burden reduction, design of a statistical production process when knowledge is explicitly considered in the business model. This approach is important to enrich the output, to strengthen the consistency of official figures as well as to reduce the burden on respondents in terms of knowledge distance. It also calls for substantial changes in the way people are trained and work is organised in National Statistical Institutions.

Keywords: quality management, quality indicators, standardisation of production process, knowledge management, business model

Title of abstract: "Show me your code, and then I will trust your figures": Towards software-agnostic open algorithms in statistical production

Jacopo Grazzini, Julien Gaffuri, Pierre Lamarche, Jean-Marc Museux
Eurostat

This contribution aims at further promoting the development and deployment of open, reproducible, reusable, verifiable, and collaborative computational resources in statistical offices regardless of the platform/software in use. Motivated by the consensus that data-driven policies should be transparent, we argue that such approach is not only necessary for the practical implementation of statistical production systems, but also essential to reinforce the quality and trust of official statistics, especially in the context of a "post-truth" society. With the advent of open source in the scientific community, as well as other disrupting technologies and software emerging from data science and similar active communities, many statistical organisations are nowadays considering to introduce new software solutions in production environments so as to benefit from the many statistical libraries, advanced algorithms and innovative developments available. Owing to legacy issues, there is however a trade-off between the risks linked to business continuity and reliability against the need for efficiency, innovation and cost-effective solutions. We devise some practical require-

ments to gear the continuous and flexible development and deployment of (open and free, but proprietary as well) software production components before any solution being adopted. Together with some best practices derived from the open source community (version control, automated unit tests, generic documentation, continuous integration, and collaborative development) we propose to unleash the social power of open algorithms so as to create new participatory models of interaction between „produsers" (statisticians, scientists and citizens) that can contribute to a more holistic and extensive approach to production systems. Overall, a greater transparency in designing production processes is expected to result in a better grip on the quality of the statistical processes involved in data-driven policy-making. We illustrate this flexible and agile approach with various open, stand-alone software or source code used in the actual production of official statistics at Eurostat.

Keywords: statistical production, open algorithms, produsers, reusability and reproducibility, control and maintenance, sharing and collaboration, software-and technology-agnostic implementation, micro services, open source and legacy software

SESSION 21 – QUALITY INDICATORS

Title of abstract: Quality indicators of the EuroGroups Register

Zsolt Völfinger
Eurostat

The EuroGroups Register (EGR) is a statistical business register focused on multinational enterprise groups active in Europe. To create the EGR, Eurostat collects input information on group members and on their relationships from the national statistical business registers of EU and participating European Free Trade Association countries and from commercial sources. After consolidation and validation, the register contains the global structure of the multinational enterprise groups. When compiling the EGR data of multiple sources has to be treated and processed. Different sources deliver complementary and partly conflicting information on the constituent units of the groups and their relationships. To monitor the EGR production and to evaluate the results of the register Eurostat developed a wide range of quality indicators for the EGR process.

The paper will present the system of the quality indicators built behind the EGR production process:

- input indicators to measure the completeness and consistency of the input data files;
- throughput indicators evaluating the data processing results in the different steps of the EGR production process;
- global output indicators measuring the completeness and consistency of EGR final data;
- micro-level output indicators comparing EGR data to official EU statistics on foreign affiliates
- definition of appropriate thresholds for the indicators;
- development of a composite indicator to measure and to report on the overall EGR quality.

The paper will also present how the EGR process is improved and which actions are planned based on the results of the EGR quality indicators. Comparison of the quality indicators for two consecutive EGR production cycles will be also presented.

Keywords: business register, EuroGroups register, process quality indicators

Title of abstract: Assessing the uncertainties of statistical outputs: the case of purchasing power parities

Laurent Olislager
Eurostat

Purchasing power parities (PPPs) are indicators of price level differences across countries. They are used for a wide range of purposes. In the Eurostat-OECD PPP exercise, prices are collected via specifically-designed surveys, for a variety of products and services covering an important subset of economic activities. For each country and item, the average price of the item is compared to the prices of comparable items in other countries. PPPs are obtained by averaging all the bilateral price ratios for products and services grouped in pre-defined categories. Using weights from National Accounts, PPPs are aggregated up to the GDP level. The methodologies of the surveys are designed to guarantee comparability and reliability of the results, which are subject to a thorough validation process incorporating various quality measurements. However, due to

the use of purposive sampling and the multilateral nature of PPPs, it is not possible to calculate precise error margins for PPPs. Experience suggests that differences between countries of over two percentage points are generally statistically significant. In this work, we aim at better characterizing the uncertainties of PPPs. We have carried out various sensitivity analyses to assess the robustness of the results when subject to changes in the input data and/or implicit assumptions in the calculation. General trends confirm the high quality of the results and can contribute to an enhanced communication with stakeholders. Our results also allow to assess the relative impacts of the input data on the final results; it can contribute to enhancing the efficiency and effectiveness of the production process, by helping in the optimization of the design and validation.

Keywords: international comparisons, price statistics, purchasing power parities, uncertainty estimation

Title of abstract: How much has national information changed in the Eurogroups Register? Some quality indicators in the scope of the ESBR Project

Luis Esteban Barbado Miguel
National Statistics Institute of Spain (INE)

The European System of Interoperable Business Registers (ESS.VIP.ESBRs) is a core Project of the VISION 2020. It aims to create a coordinated infrastructure of frames, serving as backbone for the production of high-quality business statistics in Europe. The Euro Groups Register (EGR) is the heart of the ESBRs system. The EGR contains structural information on multinational groups in the UE. It is managed by a cooperation model based on multiple integrations of microdata provided by Eurostat and the NSIs of European Member States and EFTA countries. The inherent complexity of the EGR business process and the need to reach a good level of stabilization, make the design of active quality policies needed. In fact, a specific Data Quality Program (DQP) is already in implementation phase. This context will reinforce the role of the EGR as the authoritative frame for globalization statistics in the ESS. This report takes the main elements stablished in the above mentioned DQP, but using an alternative approach. It provides some quantitative indicators reflecting the changes produced between the Spanish data provided in the EGR production cycle and the information finally registered in the global EGR frame. In addition, the relative importance of the changes produced and their accuracy are also explored, through specific desktop research. The final aim is to provide a wider statistical basis, helping to a better delineation of the priority quality actions for the EGR.

Keywords: Business Register (BR), European System of Interoperable Business Registers (ESBRs), Euro Groups Register (EGR), Data Quality Program (DQP), European Statistical System (ESS)

Title of abstract: Building Sustainable International Statistical Products: Policies, Technology, and Quality Returns from the World Development Indicators

Biokou Mathieu Djayeola
World Bank, Washington D.C.

The World Bank is a global leader in international development financing. Its World Development Indicators are the primary collection of development statistics including economic, social, and environment indicators compiled from officially-recognized international sources. It presents the most current and accurate global development data available, and includes national, regional and global estimates relevant for the purposes of analytics, operational, and policy making decisions in international development arena. In this paper, we review major policy decisions, key investments in technological platforms, and best practices in quality management around the reputation of the World Development Indicators. These findings are presented with the prospects of assessing how international statistical organizations model high demand statistical products, gather resources for their production, and establish subsequent quality frameworks – including customer relations – needed for their sustainability. We explore how specifically tailored strategic positioning contributes to the development of an ecosystem of interrelated products, setting the highest standards for producers, brokers, and users of international development statistics. Finally, we discuss the economics of quality statistics management in an evolving technological environment.

Keywords: quality management, statistical products, international development, technology, customers' relations

SESSION 22 – QUALITY OF FLASH ESTIMATES

Title of abstract: Flash estimates of income distribution indicators for the European Union: results 2016, methodology and quality assessment

Aura Leulescu, Cristina Lopez-Vilaplana
Eurostat

Indicators on poverty and income inequality represent an essential tool to monitor progress towards the Europe 2020 poverty and social exclusion target and to prepare the European Semester. They are based on EU statistics on income and living conditions (EU-SILC) and are available for all countries around 18 months after the reference period. In order to better monitor the effectiveness of social policies at EU level, it is important to have more timely indicators. A new approach was therefore proposed, which consists in the development of flash estimates (FE). FE have currently a release date approximately one year earlier than the actual data. The main methodology used is based on microsimulation techniques further enhanced to take into account the evolution in employment, population structure and indexation factors. Developing flash estimates on poverty and income inequalities in the ESS involves that their methods, sources and output adhere to a common quality framework. This includes: 1) quality checks concerning input coherence and intermediate steps 2) the historical performance of the model is defined as the ability to predict accurately the past changes as captured by EU-SILC and 3) the plausibility of the estimated changes, mainly by linking these to evolutions in observed indicators (e.g. employment trends, total household income in national accounts, national data) and disentangling the impact of simulated policies via EUROMOD. EUROSTAT has published for the first time FE 2016 as experimental statistics. While there are still limitations and we cannot expect the estimates to capture perfectly EU-SILC changes, the FE can provide useful information about the direction and magnitude of the change. The FE 2016 include several indicators, including the at-risk-of-poverty and interquartile share ratio. Deciles that measure changes at different points of the distribution seem to be important complements as early warnings for yearly changes.

Keywords: flash estimates, income, poverty, timeliness, microsimulation, policy effects

Title of abstract: ESS-cooperation on Employment Flash Estimates

Christine Gerstberger, Véronique Deneuville, Jenny Runesson, Johannes Wouters
Eurostat

Following the successful introduction of EU and euro area GDP flash estimates after t+30 days in April 2016, Eurostat and national statistical institutes (NSIs) are cooperating on a new project: the estimation of EU and euro area employment flash estimates in national accounts, which could be published at t+45 or t+30 days. A task force has been set up in 2017 to work on test estimates and quality criteria to decide on the feasibility of such estimates. The task force allows NSIs to share their experience on the use of available source data as well as estimation techniques. These will be documented in a manual to provide methodological guidance to NSIs and information to users. While the estimation of European aggregates is obtained as a weighted average of countries' growth rates, econometric techniques will be explored to complement for missing country data and/or to address a systematic bias in the estimates reflecting countries' revisions. Following the first test estimates, country coverage and results look promising for a t+45 flash estimate. Estimates

at t+30 will be explored, but are more challenging, since countries generally need to use modelling techniques to estimate missing data for the last month. The decision on the publication of EU and euro area employment flash estimates will be taken based on live estimates and back estimates and on quality criteria combining minimum country coverage and reliability of the estimate. While it is too early to draw conclusions at this stage, Eurostat is optimistic that the key macroeconomic indicators can be complemented by employment flash estimates. In any case, the work of the Task Force is an appreciated opportunity to share experience between NSIs and a good example of cooperation within the ESS to improve the availability of high quality statistics for users.

Keywords: Flash, employment, estimates, quality, criteria

Title of abstract: Flash estimate of the poverty rate using microsimulation: estimations based on French data

Kevin Schmitt, Michaël Sicsic, Laurence Rioux
INSEE, Paris, France

Every year in September N+2, INSEE publishes the poverty rate and the main indicators of inequalities in standard of living for year N. This delay is unsatisfactory for meeting the social requirements of users of these indicators. Of the 21 months between the end of the year under consideration and the publication of the poverty rate, about three-quarters of this time is taken up collecting tax and social data, and about one quarter with statistically matching Labour Force Survey (LFS) data, from which the Tax and Social Incomes Survey (ERFS) is produced. Nowcasting consists of producing an earlier indicator of the poverty rate for the target year N (in autumn N+1) based on the ERFS N-1. The method to be used here is microsimulation, which creates individuals' standard of living by imputing benefits and contributions on scales, and thus it is possible to take account of any legal changes made to these measures. The exercise is based on the INES model, which simulates the majority of French social security and tax legislation, based on any year of the ERFS. To implement nowcasting, one important step is ageing population by uprating incomes (using surveys about wages, aggregated tax data, inflation...) and calibration weighting (using margins from LFS and census). Reverse ageing is also used so that evaluations for year N and N-1 (and thus annual evolutions) are only based on the ERFS N-1 (that is minimising the sample bias). In this paper, we present the methodology and assess the quality of the early indicators thus produced. Indeed, we compare the results that would have been produced by microsimulation with those that were in fact disseminated from the ERFS. When applied to the target years 2010 to 2015, this method producedflash estimatesimilar to the actual figures published the following year.

Keywords: Flash Estimates, Nowcasting, Microsimulation, Poverty and Inequalities indicators

Title of abstract: Nowcasting Finnish Real Economic Activity: a Large Dimensional Approach

Paolo Fornaro - The Research Institute of the Finnish Economy, Finland
Henri Luomaranta - Statistics Finland

Despite the evergrowing data availability, statistical institutes publish economic indicators with considerable lag and the initial estimates are revised considerably over time. In Finland, the first estimate of GDP provided by Statistics Finland is released 45 days after the end of the reference quarter (flash estimate), while the first „appropriate" version is released 60 days after the end of the quarter. We develop a nowcasting framework in order to provide faster estimates of the monthly real economic activity indicator, the Trend Indicator of Output (TIO), and of quarterly GDP. In particular we rely on firm-level turnovers, which are available shortly after the end of the reference month, to form our set predictors. Given the large dimension of our dataset, we rely on a set of statistical models and machine learning methodologies which are able to handle high-dimensional information set to compute the nowcasts. The results of our pseudo-real-time analysis indicate that it is possible to provide substantially faster estimates of the TIO and GDP without increasing the revision error. Finally, we examine the nowcasting accuracy obtained by relying on traffic data extracted from the Finnish Transport Agency website, and find that using machine learning methods in combination with this big-data source provides competitive predictions of real economic activity indicators.

Keywords: nowcasting GDP, machine learning, model combinations, big data, timeliness

Title of abstract: Use of alternative data sources as flash estimates of economic indicators – abstract

Črt Grahonja
Statistical Office of the Republic of Slovenia

At the Statistical Office of the Republic of Slovenia we calculate the statistics of gross domestic product (GDP) every quarter of a year. 60 days after the reference period is unfortunately the quickest we can publish such statistics as the timeliness of GDP data is limited by the survey evaluations of some of the components that make up GDP. However the use of flash (rapid) estimates could fasten this process. On the basis of investigation of various big data sources we had the idea to use the data we acquired from traffic sensors and use them as primary and secondary regressors in a linear regression model for nowcasting GDP 45 days after the reference period. Nowcasting is a method of calculating estimates on the basis of unknown present or near-future values with the use of a known correlator. In the following article we describe our work and the process of nowcasting indicators from the point of data acquisition to the end results on GDP and also on a known GDP correlator, the Industry Turnover Index. We also touch on what could be extended in the future like component estimation, model accuracy improvement and data processing improvements. We wish to show how useful such data can be and what was needed to be done, before this data could actually be used. Different types of knowledge were used while composing the selected process for economic indicator estimation. These include new skills in the fields of information technology and methodology, and knowledges in the respective subject matters.

Keywords: nowcasting, big data, linear regression, GDP, economic indicators

SESSION 23 – APPROACHES FOR MULTI-SOURCE STATISTICS

Title of abstract: Enhancing the quality of National Accounts in the estimation of proceeds from illegal markets

Tiziana Tuoto, Antonella Baldassarini, Valentina Chiariello
Italian National Institute of Statistics (ISTAT)

National Institutes of Statistics need to assess the size of some illegal activities considered as a substantial part of the illegal economy in order to understand their value and impact for society. At European level, it has come to the decision to include illegal activities that produce goods and services in National Accounts. The accounting data provided to the European authorities by each country must include the income generated by the markets of drugs, prostitution and the smuggling of cigarettes and alcohol. Currently, in the GDP, several European National Institute of Statistics estimate the proceeds of illegal drug markets on the demand side (data on the number of consumers and quantities consumed), prostitution on the supply side (data on the number of prostitutes on the market) and cigarette smuggling also on the supply side (data on seizures). In this work, we aim at providing an accurate estimate of the size of these illegal markets. Moreover, we propose a uniform procedure for measuring the flow of illegal proceeds in the GDP. The estimation methodology applies in steps: first we calculate the hidden population of illegal workers in illegal markets of drug, prostitution and smuggling to give a more precise dimension to these markets. To this purpose, we exploit administrative data coming from the Ministry of Justice, recording criminal charges of Public Prosecutor's offices. Due to privacy motivations, we only have soft personal information about the criminals. However, we can consider the administrative source as a list of criminals and by this personal information we are able to count how many times a criminal appears in the list. The main question is how many criminals are missed by the justice system but active in the illegal markets and we provide methodologies to answer to this point.

Keywords: illegal markets, hidden population size estimation, data integration, national accounts

Title of abstract: Integration of inconsistent data sources using Hidden Markov Models

Paulina Pankowska, Dimitris Pavlopoulos - Vrije Universiteit Amsterdam, Netherlands
Bart Bakker - Statistics Netherlands
Daniel Oberski - Utrecht University, Netherlands

Latent class models (LCM) are increasingly used to estimate and correct for classification error in categorical register and survey data, without the need for a "gold standard", error-free data source. To accomplish this, LCMs require multiple measures of the same phenomenon within one data collection wave ("latent structure model"), or over time ("hidden Markov model"), and assume that the errors in these measures are conditionally independent. Unfortunately, this "local independence" assumption is often unrealistic, untestable, and a source of serious bias. However, linking independent sources can solve this problem by making the assumption plausible across sources, while potentially allowing for local dependence within sources. Thus, while an attractive method for the production of more accurate official statistics, this procedure is very complex and time consuming. More specifically, the use of LCM often requires performing linkage

between different data sources and re-estimating the model for each new time period. What is more, data linkage might lead to linkage error and subsequently to biased estimates. In our research we investigate the feasibility of using HMMs specifically in the production of labor mobility estimates in the Netherlands using register and survey data. We do so by first looking at the possibility of parameter re-use and then by analyzing model sensitivity to linkage error. The results suggest that the HMMs error estimates are time invariant and, therefore, can be re-used in later time points without the need to re-link the datasets and re-estimate the statistical model. The results also show that linkage error only leads to (substantial) bias in very extreme scenarios and that HMMs, to an extent, can correct for false-positive linkage error. It would be very nice to let this presentation be preceded by Bakker and Zult et al. in the same session. The first paper introduces the topic.

Keywords: Measurement error, Hidden Markov Models, Latent class modelling, linkage error, data linkage

Title of abstract: On the Experimental Usage of Ontology-based data access for the Italian Integrated System of Statistical Registers: Quality Issues

Raffaella Aracri, Adele Bianco, Roberta Radini, Monica Scannapieco, Laura Tosco - Italian National Institute of Statistics (ISTAT)
Federico Croce, Maurizio Lenzerini, Domenico Fabio Savo - Sapienza University of Rome, Italy

Ontology-based data access (OBDA) is a recent paradigm for addressing data management based on a conceptualization of the domain of interest, called ontology. A system realizing the vision of OBDA is constituted by three layers: the ontology, that provides a high level, formal, logic-based representation of the above mentioned conceptualization; the data source layer, representing the existing data in the various assets of the system; the mapping between the two layers, which is an explicit representation of the relationship between the data sources and the ontology. Most works on OBDA focus on querying data through the ontology. However, recent papers argue that OBDA is a promising tool for assessing the quality of data, especially in the presence of multiple, possibly mutually incoherent data source. For example, with the OBDA approach, checking data consistency reduces to verify whether the data sources contain data contradicting the axioms constituting the ontology. We have experimented the above approach for a current project of Istat, namely the Italian Integrated System of Statistical Registers. We have focused on the domain of population data, and we have built an OWL ontology for modeling basic concepts and relationships of this domain, including persons, families, parental relations, citizenship, locations, etc.. Then, we have considered a core set of population data and we have specified the mappings from such data sets and the ontology. With such a specification at hand, we have used the MASTRO system for OBDA for carrying out several data quality checks, focusing in particular on the consistency dimension. The preliminary results are extremely encouraging, both in terms of effectiveness of the method and in terms of efficiency of the checking procedures, in the sense that the performance of the quality check is not affected by the (usually expensive) task of reasoning over the ontology.

Keywords: ontology, data integration, statistical register, data quality

Title of abstract: Data quality and data integration

Raffaella Traverso, Bernadette Lauro
European Central Bank, Frankfurt, Germany

Data integration allows users to have a unified view of data coming from different data sources. However, data sources can be characterized by several types of heterogeneities such as different data models, different data representations and different quality at the level of data instances. As the widely accepted definition of Data Quality is "fitness for purpose", we pose the question: how is the quality of the source data affected when data are combined to satisfy the integration purpose? In the process of integrating data, problems in the quality of data sources become more evident. These are measured by relevance, accuracy, reliability, timeliness, consistency of the data as well as by the completeness, validation and relevance of the structure of the metadata attached to the data sources. Through the use of practical examples we present the challenges posed by data integration and how data sources are thereby affected.

Keywords: Data integration, data quality

SESSION 24 – DATA INTEGRATION METHODS AND ARCHITECTURES

Title of abstract: Improvement of the capture-recapture model to estimate under coverage of administrative data sources

Daan Zult
Statistics Netherlands

The under coverage of an administrative data source is the true population size minus the observed population size. The true population size is often estimated with the capture-recapture (CRC) model. An important requirement for the CRC model to hold is that specimens over different captures can be identified such that the observer knows whether they are the same or not (i.e. no linkage errors). This is particularly relevant when identification is not obtained by some unique identifier (like a tag or id-code) but by background information (like name, address or skin patterns). Linkage based on background information is usually probabilistic and can lead to different specimens being falsely linked or the same specimen being missed as link, which gives biased population estimates. A partial solution to this problem was provided by Ding and Fienberg (1994) and was later extended by Di Consiglio and Tuoto (2015). These authors show how to use linkage probabilities to correct the CRC estimator. De Wolf et al. (in progress) further extend this model and shows that all these models are special cases of a more general model. However, all these correction methods are designed for only two registers and no use of background variables (such age sex or age), which implies that correcting for linkage errors cannot be combined with other corrections that are often required to correct for other sources of bias. In this paper we reformulate the general model of de Wolf et al. into a Weighted-CRC (W-CRC) model. W-CRC can deal with multiple captures and background characteristics. We further show how the W-CRC model works in practice.

Keywords: Capture-recapture, Log-linear models, linkage errors

Title of abstract: Change is the law of life. Prepare your data warehouse for a Big Future, by including Big Data

Antonio laureti Palma, Gianpiero Bianchi - Italian National Institute of Statistics (ISTAT)
Sonia Quaresma - National Statistics Institute of Spain (INE)

National Statistical Institutes, NSIs, produce statistics based on consolidated statistical models according to the study-domain characteristics. Conversely, nowadays, Big Data (BD) opportunities force NSIs to deliver statistical products based on a sequence of data analytic processing. This feature means that algorithm experts may abstain from being informed theoretically about the subject matter; i.e. in data analytics we can say that knowledge extraction is data driven. From this perspective, NSIs multidisciplinary approaches are crucial. This is because they must be able to combine different areas of expertise for the mapping of analytical constructs with theoretical concepts. This is an epistemological process change that forces NSIs to meet some basic requirements at the procedural, organizational and infrastructural levels. At the process level, the use of BD introduces a paradigm shift, typical of data mining, from theory to data driven models. At the Organizational level, it is necessary to guarantee and support the active participation of multidisci-

plinary experts in the process of knowledge extraction. At the infrastructural level, new and complex infrastructures are needed to support both analytical tools and multidimensional analysis. Typically, in a NSI the central repository of integrated data is the corporate Statistical Data Warehouse (S-DWH); this stores current and historical data in one single place and it is used by knowledge workers. To manage the opportunities arising from BD the S-DWH must address the needs of new data types, new volumes, new data-quality levels, new performance, new metadata, and new user requirements. The S-DWH will be data-driven and extremely flexible and scalable. In this work we aim to present some S-DWH architectures involving integration with the typical prospective BD as possible statistical sources: sensor data, webscraped data, mobile phone data. All the solutions presented aim to facilitate a multidisciplinary approach to improve statistical process quality.

Keywords: data integration architecture; official statistic production; statistical data warehouse; big data; data mining architecture; data analytics architecture; data driven

Title of abstract: Combining information from different sources to estimate the annual working hours of part-time workers

Sandra Barragán Andrés, Elisa Martín Hernández
National Statistics Institute of Spain (INE)

Structural statistics on earnings provide comparable information on relationships between the level of earnings, individual characteristics of employees (sex, age, occupation, length of service) and their employer (economic activity, size of the enterprise). In Spain there is a main survey carried on 4-year periodicity. Additionally, due to administrative records, most of the information can be given annually as well. However, some problems arise when different administrative records are used together and joined with survey results to obtain the final dataset. Administrative data were originally collected for a definite non-statistical purpose that might affect the quality of the data. In addition, there is a coherence issue when more than one source of information are integrated to obtain data for the same variable. Coherence is an essential part of official statistics so it is receiving such a growing attention to ensure quality in the data. Statistical offices focus on the importance of obtaining coherent results specially when they come from a process of data integration. Along this process, we deal with the inconsistency of the individual values by taking into consideration the additional uncertainty due to the difference between the administrative concept and the statistical variable. Therefore, we present the methodology used in the last publication of the structural earnings survey to obtain the annual working hours of part-time workers. We have combined three sources of information in order to obtain a trustworthy value of the working hours for each individual. We have developed an algorithm with decision rules built by using expertise on the field. This algorithm is an initial solution to maintain coherence at microdata level in the combination of different sources for the same variable.

Keywords: Data integration, Administrative records, Combining information, labour market, working hours

Title of abstract: Data Architecture for Statistics Production – Logical Data Repositories

Tuukka Saranpää
Statistics Finland

Traditionally, data repositories at Statistics Finland have been designed for the need of one or a few specific statistical areas. Additionally, development projects on data repositories have typically assumed a systems-oriented approach, focusing on the technological solutions required. Over time, the lack of top-down coordination has resulted in overlap, inconsistency, poor opacity and impaired quality between different data repositories. These issues, along with new national guidelines for Finnish public administration, prompted Statistics Finland to organize a preliminary study to outline the fundamental requirements for a common statistical data architecture as well as a roadmap for its development. In January 2017, following the findings of the preliminary study, Statistics Finland launched a development program on data architecture for statistical production to coordinate the various projects outlined in the roadmap. In order to ensure consistency and interoperability between different development projects, a top-level to-be model for logical data repositories for statistics production was required. In developing the model, Statistics Finland assumed a holistic top-down view on statistical production, shifting the focus from statistics to data. The model envelops the entire statistics production process (Generic Statistical Business Process Model phases 4-7,) from data acquisition to dissemination and determines the names and general constitution of all relevant top-level data repositories. Some key observations of the new model are the emphasized role of metadata through the process and a centralized top-level repository for geospatial data serving all areas of statistics production. The newly-developed model provides a top-level framework for any development effort on a specific area. However, to truly capture the benefits of a common data architecture, significant bottom-up development work is required to determine the exact content and physical infrastructure of each repository. Furthermore, the top-level model may be improved based on observations and experiences from lower-level development.

Keywords: Data architecture, Statistics Production, Logical Data Repositories

Title of abstract: Modernizing Data Integration Systems at Istat

Piero Demetrio Falorsi, Roberta Radini, Monica Scannapieco, Laura Tosco
Italian National Institute of Statistics (ISTAT)

Istat has engaged a modernization programme that includes a significant revision of the statistical production. The principal concept underlying such an important change is the usage of a system of integrated statistical registers as a base for all the production surveys; this system will be in the following referred to as the Italian Integrated System of Statistical Registers (ISSR). The ongoing work for building such a system required a big investment on architectural aspects to guide the enterprise-level design of the system. In this respect, there are two major activities that have been engaged: (i) the design of the data architecture of the ISSR and (ii) the design of the processes to populate the ISSR. The data architecture of the ISSR has been conceived according to modern semantic integration approaches, with a strong emphasis on active metadata guiding the access to the system. The design of the processes has been performed by relying on existing standards, like GSBPM and GAMSO, and frameworks, like ESS Enterprise Architecture Reference Framework. In this paper, we will describe both strands of work, focusing on how quality aspects have been taken into account and specifically on the role played by both official statistics and world-wide technical standards to ensure the quality of data, processes and underlying information systems.

Keywords: data integration, process modelling, data architecture

SESSION 25 – DATA INTEGRATION AT WORK

Title of abstract: The use of loyalty card transaction data from retailers in Household Budget Statistics: an exploration

Sverre Amdam, Bart Bulenes, Henning Holgersen, Dag F Gravem
Statistics Norway

Statistics Norway (SSB) is exploring different alternative data sources for household budget statistics. One such source is the use of loyalty membership data from retail chains, where particularly data from grocery retail chains (food and drink consumption) is of interest. The paper presents an explorative analysis on the use of this type of data from a methodological perspective, using a test-sample of receipt data with information on loyalty members from one of the largest grocery retail chains in Norway, containing over 121 million transactions, about half of them conducted by members. The main finding in the analysis is that using loyalty members purchase data from one retail chain to estimate Norwegian households grocery consumption is not advisable, even though data derives from a large retail chain with high market share. Members have a spending pattern that is different from non-members, and further different from the general population. Not only are the members a specific subpopulation when it comes to spending, there are also demographic differences. Considering gender and age, the members are a selective subset, and we find that differences in spending varies in different age groups and between men and women. The observed skewness of the members with respect to background variables, such as gender and age, suggests that an adjustment for these variables may remove some of the selection bias seen in the expenditure pattern. We found that this is not the case. Adjusting for age and gender only has a very small effect on the resulting spending pattern. We conclude that the members are a specific subpopulation with a spending pattern unlike that of the average Norwegian, and that this cannot be explained by age and gender. There must be other factors at play that are not available in the data used in this analysis.

Keywords: HBS, big data, transaction data, loyalty card data, alternative data sources

Title of abstract: Smart business statistics: how to integrate technology and official statistics?

Ger Snijkers, Sofie De Broe
Statistics Netherlands

Businesses are requested to deliver a lot of data for official statistical purposes. Traditionally, questionnaires are used as collection instrument, which represents high costs and response burden for businesses. Nowadays, however, a lot of data within businesses are already available in electronic format. Electronically available data could be collected by means of System-to-System (S2S) data-communication technology as a primary data-collection mode. An example is financial data in a computerized business information chain which automatically provides input for financial, tax and statistical reports. These data can be collected using e.g. SBR (Standard Business Reporting), a standard for S2S data communication. A second example regarding S2S data collection for statistical purposes involves sensor data. Increasingly electronic sensors are used to run a business, e.g. by agricultural (like dairy farms) and transportation businesses. In this paper we explore the challenges and consequences of S2S data collection. Drivers for switching from questionnaires

to S2S data collection are working towards smart business statistics, (i.e. timely and new statistical output integrated in business processes), the reduction of response burden, and monitoring and benchmarking businesses to their counterparts. S2S data collection seems straightforward. However, important factors that affect or impede implementation are standardisation of S2S technology and harmonisation of metadata. Also definitional issues and trust in the quality of the data affect its use. S2S data collection offers opportunities to get access to data in a cost-effective manner (as opposed to sending out questionnaires); however, both companies and NSIs may need to do initial investments in the technology. In terms of getting access to the data and business participation, the question from businesses remains: „what's in it for me?"

Our statement is: Technology is the enabler of innovations; it is the applied methodology and organizational context that make innovations work.

Keywords: automated data exchange, electronic data interchange (EDI), sensor data, data quality, enablers and challenges

Title of abstract: Utilizing Non-Probability Survey Data to Improve the Quality of Probability Survey Data for Small Samples

Joseph W. Sakshaug - Institute for Employment Research and University of Manchester, United Kingdom
Arkadiusz Wisniowski, Diego Perez-Ruiz - University of Manchester, United Kingdom
Annelies Blom - University of Mannheim, Germany

Probability-based surveys serve a useful role in society as they provide a critical source of representative data used to track changes in the population over time, evaluate the impact of interventions, and inform policy-relevant decisions. However, probability-based surveys are struggling against increasing non-response rates and non-coverage rates in western societies, which have contributed to rising costs of data collection that have outpaced research budgets. To cut costs, surveys may choose (or be forced) to reduce sample sizes, but this comes with the drawback of increased variability of survey estimates and a higher likelihood of Type II errors. In this paper, we evaluate a Bayesian modeling approach for reducing the variability of survey estimates derived from small probability samples. The approach incorporates auxiliary information collected from concurrent (and larger) non-probability samples into the estimation process. We demonstrate the method using actual nationally-representative probability and non-probability survey data collected in Germany. We show that the data combination method produces survey estimates with substantially lower variability compared to probability-only survey estimates, especially for small sample sizes. We conclude with a discussion of the implications of this procedure for survey practice and propose some future research directions.

Keywords: Bayesian modelling, data integration, estimation, small sample inference

Title of abstract: Design and evaluation of an editing and imputation strategy for micro-data from integrated administrative sources: the Italian case of the ARCHIMEDE Project

Romina Filippini, Sara Casacci
Italian National Institute of Statistics (ISTAT)

In order to provide users with detailed statistical information at local level, in recent years the Italian National Institute of Statistics (Istat) has made available micro data collections based on the integration of several administrative sources. In particular, micro data archives have been produced in the context of the ARCHIMEDE project (Integrated Archives of Economic and Demographic Microdata) to enlarge the offer of statistical information on the households socio-economic conditions. Extracting reliable statistical information from multiple data sources is in general a complex task. The present work describes the methodology adopted to ensure more „complete" and „coherent" data dissemination and provides indications regarding the quality of the results produced. Specifically, the different editing and imputation techniques and tools used for the main variables in the database are illustrated. A specific focus concerns the consistency between the two quantitative variables "Income from employment" (from fiscal sources) and "Work Intensity", (from social security data). The latter, taking values in [0,1], is defined in terms of amount of work carried out over 1 year. Before applying the error localization and imputation procedures, a detailed analysis based on auxiliary administrative sources has been conducted aimed at properly identifying erroneous cases to be corrected. Finally, impact of imputation at territorial level is evaluated by comparing values of some indicators based on the "raw" data, with the corresponding values based on the final adjusted data.

Keywords: editing and imputation, administrative data, experimental statistics

Title of abstract: Reconciling Estimates of Demographic Stocks and Flows through Balancing Methods

Marco Di Zio, Marco Fortini, Diego Zardetto
Italian National Institute of Statistics (ISTAT)

In the near future, the Italian population census will be the result of the integration of administrative and survey data. This will enable Istat to deliver official population size estimates more frequently than it happened before through traditional censuses. Census-based estimates of population counts ('stocks') should be consistent with information about demographic events ('flows') available from municipal civil registries. In particular, the Demographic Balancing Equation (DBE) should be fulfilled, which states that final population counts P(t+1) are equal to starting population counts P(t) plus the sum of natural increase N (difference between births and deaths) and net migration M (difference between immigrants and emigrants):

Due to sampling and non-sampling errors, the DBE will not be trivially satisfied. Therefore, suitable methods must be investigated to obtain consistent final estimates. These methods should simultaneously adjust both the initial estimates of population counts and the rough civil registry figures, in such a way that the resulting macrodata exactly fulfill the DBE.

We formalize the problem of ensuring time and space consistency of demographic estimates as a constrained optimization problem. Given initial, rough estimates of stocks and flows entering the demographic balancing equations defined for all the geographic areas of a given territorial level, we search for final estimates that arebalanced, i.e. (i)satisfy all the DBEs, and (ii)are as close as possible to the initial estimates. To solve the problem, we propose to exploit the Stone-Byron approach that is commonly adopted for balancing large systems of national accounts.

Experiments on real data suggest that, under reasonable assumptions, the proposed approach determines improved estimates of population counts: besides gaining consistency, they exhibit lower bias and variance as compared to rough ones, and the observed efficiency gain seems robust against misspecification of reliability weights.

Keywords: Macro-integration, demographic balancing equation, data integration, administrative data, census estimates

SESSION 26 – INNOVATION IN DATA COLLECTION I

Title of abstract: Geospatial mobile data to increase the quality of usual place of residence

Kaja Sõstra, Kristi Lehto
Statistics Estonia

The next population and housing census in Estonia at the end of 2020 is intended to be register-based. One of the biggest challenges on the way to register based census is the difference between registered and actual places of residence. By law, everybody is obliged to ensure that their correct usual residential address is entered in the Population Register, but there are different reasons why people don't do that (for example school places, different benefits etc). The difference between registered and actual places of residence affects the breakdown of the lowest level of the place of usual residence and all household and family characteristics. To solve this problem Statistics Estonia (SE) has started pilot project for testing possibility to use mobile positioning data. The pilot study consists of four steps. Firstly SE asked volunteers to participate in the project and to fill in written consent to use their mobile positioning data. In addition the actual place of residence of volunteers was collected. Secondly the set of potential addresses was created for each participant by linking the addresses of respondent and his/her close relatives from Population Register and Land Register. The anchor points of the first and second place of residence are estimated based on the mobile positioning data as a third step. Finally anchor points and other auxiliary information is used to build a model for selecting the most probable place of residence from the set of addresses. Results of the pilot study will be presented.

Keywords: register based census, mobile positioning, anchor point model

Title of abstract: Mixed Mode Effects of Web and Telephone Surveys on Measuring Employment Status

Joachim Schork, Cesare Riillo, Johann Neumayr
STATEC, Luxembourg

Web questionnaires are increasingly used to complement traditional data collection, leading to different combinations of mixed survey modes. The flexibility of mixed mode surveys provides many advantages such as broader coverage of the population, less nonresponse issues, lowered expenditures, and compensation for the decreasing availability of other data sources, i.e. fixed-line telephone numbers. However, the increased usage of web data raises concerns whether web questionnaires lead to systematic measurement errors, since responses given to web questionnaires may be significantly different compared to other survey modes.

Based on the Luxembourgish Labor Force Survey, the largest recurring household survey in Luxembourg, we investigate if reports on participants' employment status differ in web and telephone data. Analysis of the raw data reveals significant differences in sample composition (e.g. respondents' personal characteristics such as age or nationality) as well as in key variables for measuring employment status. People in the web sample report, for instance, more often to be in employment, work more hours on average, and have more often a full-time job.

In order to investigate whether differences in employment status are caused by sample composition or survey mode effects, we match web and telephone samples according to variables that lead to dissimilarities in sample composition. We identify these variables by a combination of automatic variable selection via random forest and a theory driven selection. Based on the selected variables, we then apply a Coarsened Exact Matching to enable a comparison among web and telephone samples.

After matching, we show that most differences between web and telephone samples disappear, which leads to the conclusion that survey mode has a minor effect on measuring employment status. The results of the present study therefore support the implementation of mixed survey modes in labour market related official statistics such as the Labour Force Survey.

Keywords: web survey; telephone survey; mode effects; Coarsened Exact Matching; Labour Force Survey

Title of abstract: Innovation in questionnaire design: the electronic questionnaire 'authoring' tool

Jordan Stewart, Daisy Hamer
Office for National Statistics, Newport, United Kingdom

The Office for National Statistics is committed to moving all business surveys online in the next few years, as well as making the Census online mode the default method of data collection. So far ONS has moved five business surveys online with a high degree of success (which has been evidenced by quicker returns and fewer errors when compared to the paper versions); however the process for moving these surveys has been slow. Previously, designing the online version of paper questionnaires was very time consuming; methodologists would create the initial questionnaire templates using Google Slides while seeking feedback from subject matter experts. The agreed content would then be passed to user experience designers and developers who were responsible for highlighting problems with the design of certain questions. For example, designs might not be suitable for users who need to use assistive technology, such as screenreaders or zoom text, to help them complete the questionnaire. Subsequent redesigns were sometimes necessary. The questionnaire would then undergo pre-testing before being dispatched to respondents. In order to speed up the design and testing process an 'authoring' tool has been developed, this enables methodologists to quickly design iterations of questionnaires in the same format and style as they appear in the final version. The tool has in-built question types and answer options, such as checkboxes, radio buttons and currency values. It allows the user to add routing and validation to the questionnaire as they are building it which significantly speeds up the design process. The tool is also being developed in conjunction with a question bank, providing a unique opportunity to harmonise business survey questions, ensuring quality and consistency across business surveys. This paper will address the benefits and challenges of using the tool and discuss the future of questionnaire design in ONS.

Keywords: Innovation, Design, Quality, Modernisation, Streamlined methodologies

Title of abstract: Transition to WEB data collection in household surveys at SURS – what we have learned so far

Eva Belak, Marta Arnež, Mihaela Nemec, Mateja Zgonec, Tina Steenvoorden
Statistical Office of the Republic of Slovenia

At the Statistical Office of the Republic of Slovenia (hereinafter SURS) 20 surveys of households or individuals are conducted on a regular basis. More than 100,000 persons or households are surveyed every year (5% of the population of persons and more than 10% of the population of households). The predominant data collection mode is still telephone. Due to huge technological development of the ICT in the last 10 years and increased demand for internet-based services, and also pressure to reduce data collection costs, SURS launched a 13-month project in 2014 to set up processes and standards for WEB data collection for surveys of households and individuals. Since 2016 at SURS several household surveys have been transferred to the WEB survey mode: Consumer Survey, Mobility Survey, ICT Usage in Households and by Individuals Survey, and Household Energy Consumption Survey. The article aims to present some of the methodological challenges when transferring data collection to a new survey mode: questionnaire testing for WEB surveys, re-interviewing to possibly assess the measurement error, communication with the respondents who are the non-respondents in the presented WEB surveys, analysis of the costs, and impact of the WEB survey data collection on some key statistics.

Keywords: mixed mode design, web surveys, re-interviewing, paradata, measurement error, questionnaire testing, fieldwork strategy, costs

SESSION 27 – INNOVATION IN DATA COLLECTION II

Title of abstract: Trusted Smart Statistics: A reflection on the future of (Official) Statistics

Konstantinos Giannakouris, Fernando Reis, Michail Skaliotis, Albrecht Wirthmann
Eurostat

The extended use of the Internet of Things (IoT) will eventually take big data to a whole new level and change the data landscape. Data capturing and processing capabilities coupled with analytical and statistical capabilities will be embedded in the smart systems themselves. Intelligence along the data life-cycle enhanced with cognitive processes will be essential components of future statistics. Algorithms will handle huge amounts of data at the limit of human capabilities for exploiting statistical data using traditional data processing methods. We call this smart statistics. We identify smart statistics as the future role of official statistics in a world impregnated with smart technologies. Smart technologies involve real-time, automated, interactive technologies that optimize the physical operation of appliances and consumer devices. Statistics themselves would then be transformed into a smart technology embedded in smart systems. However, statistics are only useful when they are trusted. In order to build trust into smart statistics the data life--cycle needs to be auditable, transparent, with guarantees of accuracy and privacy by design. This paper provides a reflection on the future of official statistics in a hyper-connected world dominated by the IoT. It briefly outlines the concept of smart technologies shaping the future of statistics emphasising the need to embed trust in smart statistics under principles for transposing algorithmic transparency and accountability in smart statistics.

Keywords: Smart Statistics, Internet of Things, Smart technologies, Big Data

Title of abstract: Improving the quality of official statistics with geographical disaggregation and dasymetric mapping: Two Eurostat experiments on tourism and population statistics

Julien Gaffuri
Eurostat - GISCO

Official statistics are often reported on statistical units, which are sometimes too large to depict properly the geographical distribution of the underlying phenomenon. In the European context for example, most statistics are produced only at national level (NUTS 0) and do not allow a true understanding of the spatial pattern at more local scales. Geographic resolution is a crucial component of quality in official statistics which should be better addressed.

This article describes two experiments carried out at Eurostat for disaggregating statistics with auxiliary geographic data. These experiments are both based on dasymetric mapping: Input statistical values are distributed at the level of geographical features; these new statistical values are then re-aggregated at the level of target statistical units with a finer resolution. A first experiment was the disaggregation of tourism statistics over Europe from NUTS 2 to NUTS 3 and a 10km resolution grid. The auxiliary geographic information used is a database containing the location of around 160000 touristic accomodations over Europe.

The outcome reveals a striking image of touristic activity over Europe, with spatial patterns which cannot be revealed at NUTS 2 level. The second experiment was on the disaggregation of mobile phone data over Belgium to assess population distribution on a 1km resolution grid. Mobile phone data are collected at antenna level, whose reception zones are extremely irregular in shape and size, especially in rural areas. Cadastral information on the location and volume of each single building over Belgium has been used to locate more precisely mobile phone users around built-up areas. Both experiments show the pertinence of using geographic information with dasymetric mapping method to improve quality related to geographical resolution. This method has been implemented in the generic library EuroGeoStat (github.com/eurostat/EuroGeoStat) and is intended be applied to other domains.

Keywords: Geographic disaggregation, geographic resolution, scale, dasymetric mapping, geographic information, tourism statistics, population statistics, mobile phone data, Eurostat

Title of abstract: Investigation of linked open data technologies for publishing georeferenced statistical data

Mirosław Migacz
Statistics Poland

Polish official statistics possesses a vast amount of statistical data dispersed among different databases and disseminated using various publication methods. While there is a significant increase in openness of the data, there is still a lot of work to be done in terms of integrating different data sources. That is why Statistics Poland decided to look into the linked open data technology and launched a project for a pilot implementation based on statistical and geospatial data samples. Multiple data sources published by official statistics have been identified, described with metadata and assessed in terms of their openness. At the same time, units of territorial division of the country that are used for statistical data dissemination have been catalogued, harmonized and generalized for years 2002-2016. Finally, linked open data technologies have been explored in order to find a feasible implementation method. This pilot covered statistical data from three major databases (Local Data Bank, STRATEG system and Demography database), geographical data for statistical units and the data sources catalogue. A thorough research has been performed on existing vocabularies and statistical linked open data implementations in order to create RDF metadata and establish a test SPARQL endpoint. The pilot linked open data implementation was a valuable exercise which provided a lot of answers but at the same time raised a lot of new questions: Is there a reference implementation for statistical data? Which vocabularies to use? What should we link to? How to encode geospatial data to make them most usable? Most implementations are technically correct but are they of good quality? Hopefully an increasing interest in linked open data along with pan-European cooperation fuelled by Eurostat's DIGICOM project will provide answers to these questions and a reference statistical linked data implementation will surface soon.

Keywords: open data, linked open data, geospatial, RDF, SPARQL

Title of abstract: The impact of a centralized data collection approach on response rates of economic surveys and data quality. The Istat experience

Giampaola Bellini, Francesca Monetti, Pasquale Papa
Italian National Institute of Statistics (ISTAT)

In April 2016 Istat (Italian National Statistical Institute) started a corporate restructuring process that interested all the statistical production structures and that led to a completely renewed organizational model. Before the above mentioned reorganization, the statistical processes were organized according to the classical 'stovepipe' model, that involved independent, non-integrated, statistical processes including all the necessary skills: statisticians, information technology experts, thematic experts, methodologists. The new model restricts the thematic production processes only to the thematic experts, while all the "cross" expertise are all assigned to specialized structures. The main advantage of the new setup concerns the overall system efficiency, while the main disadvantage concern the increased fragmentation of the production processes. The new model was based on the following criteria:

- standardization and generalization of each phase of the productive process;
- specialization of personnel devoted to specific activities;
- detailed planning of the surveys calendar and of each data collection activity;
- realization a Business portal involving a set of services oriented to respondents.

Introduction of a centralized Contact center inbound and outbound Before the restructuring process, response rates in economic structural surveys (SBS, Prodcom, ICT, R&D, Innovation, Inwards and Outwards surveys) were quite low and unsatisfactory. The medium response rate was 59.9 per cent. After a year from the introduction of the new organization the medium response rate increased to 66.0 per cent. At the same time, the duration of the data collection periods reduced from 139.3 to 116.9 days. For short-term surveys the main result obtained is the increase in the number of the questionnaires transmitted by respondents within the 'useful term', as to say the deadline for the calculation of the provisional index. The paper will describe in detail the process innovations introduced and the main results achieved.

Keywords: data collection, response rates, process efficiency

Title of abstract: Mode-effects in mixed-mode surveys: the Italian experience on social surveys using the web

Orietta Luzi
Italian National Institute of Statistics (ISTAT)

Due to the ever-increasing penetration of the internet into the Italian population, Istat is progressively expanding the use of multi-mode data collection involving Computer assisted web interview (Cawi) in social surveys. In 2017 a project was carried out on the optimal design of mixed-mode strategies to ensure high data quality levels by preventing and treating mixed-mode effects. The results of the project fed the handbook „Mode-effects in mixed-mode surveys - Theoretical issues and experimental applications on social surveys using the web", which is expected to contribute to the standardization of the current social surveys. Actually, the handbook provides a conceptual reference framework in the area of mixed-mode surveys, as well as an overview of the main issues related to the design of multi-mode data collection strategies and possible methodological approaches to face them. A selection of methods for the diagnosis and the treatment of selection and measurement effects are thoroughly analyzed and assessed based on experimental applications to multi-mode social surveys. The handbook is expected to make Istat researchers more aware of quality issues related to mixed-mode data collection using the web and of the need of preventing mixed--mode effects to ensure low biasing effects on estimates. The paper summaries the theoretical contents of the handbook, focusing on the main findings and recommendations emerged from the analyses and studies made.

Keywords: Multi-mode data collection, mixed-mode effects, social surveys, Web

SESSION 28 – DATA COLLECTION, METADATA

Title of abstract: The French statistical metadata repository, RMéS: managing metadata throughout the whole statistical process

Marie Hélène Kérouanton
INSEE, Paris, France

Insee has had a long experience in metadata. Twenty years ago, Insee developed a software application for a system of structured documentation called DDS (Data structured documentation). Nowadays, DDS is quite outdated. Furthermore, the patterns used to structure metadata have not been stringent enough and don't conform to any standard. And finally, all the documentary elements are actually collected at each phase of the process, via different files. They are entered at the end of the different steps into different instances of the DDS, generating redundancies and inconsistencies. Therefore, a new statistical metadata repository, called RMéS, is being set up. RMéS relies on two repositories. The one dedicated to the questionnaires, variables and their codification is called a Colectica Repository. Information is stored in an international format, the DDI format (Data Documentation Initiative). The other one hosts all the other metadata, described in more appropriate models. These models are compliant with W3C standards and with the Single Integrated Metadata Structure as required by Eurostat for the quality reports. In this repository we store not only concepts and classifications but also the description of all our statistical sources and their quality report in a RDF format. These two repositories are linked together. At the same time, we are developing management interfaces and services to enable other applications to use the metadata. We also created an application to design questionnaires, called Pogues. With this project, we aim to gather and share in a single application all the metadata in accordance with international standards. What is more innovative is the idea of using metadata throughout the whole process, from the analysis of the needs to the statistical results and the assessment, in order to develop metadata driven processes.

Keywords: metadata, repository, standards, quality, management system

Title of abstract: Development of a statistical product portfolio tool

Rasmus Anker Kristiansen
Statistics Denmark

A regularly asked question is; how many statistical products does Statistics Denmark produce? To answer this question, it is necessary to ask, what is a statistical product? One answer could be that a statistical product is a qualitative measurement of some phenomenon. A less philosophical answer could be that a statistical product, from Statistics Denmark that is, consists of a name, one or more tables in the StatBank, a quality report with documentation, a press release, a page on the website, a person responsible and an administrative placement in the organization. Statistics Denmark is currently creating a complete product portfolio tool, which provides this overview of all our statistical products. Statistics Denmark has in recent years been working systematically with coherent metadata, including adoption of the Single Integrated Metadata Structure (SIMS). The idea is write once, use everywhere, but this raises the question; use where? The portfolio tool links all information about a given statistical product, in terms of naming, allocation in the subject structure, across dissemination platforms and administrative placement within the organization.

The portfolio tool will retrieve real time information from existing administrative systems such as the meta-data repository (Colectica), from different dissemination-, HR- and financial management systems and from the StatBank Application Programming Interface (API). There are multiple benefits and possibilities of such a tool. We get a better overview of our inventory, so it is clearer to ourselves and our users, what we actually produce. We can raise credibility when what we say we do in the annual work plan, aligns with what we actually do. We ensure coherent and consistent naming of the individual statistics and we can improve our management accounting procedures with a more precise activity-based costing (ABC) in the production of our statistical products.

Keywords: Statistical Product overview, Product portfolio, Coherent Metadata, Business Administration

Title of abstract: Labour market characteristics of European citizens living in another European country than their own (emigration perspective)

Georgiana Ivan, Mihaela Agafitei
Eurostat

Migration flows within the European labour market tend to go in specific geographic directions; some EU Member States can mostly be defined as sending countries while the others mainly as receiving ones. While the statistical characteristics of immigrant populations and their performance on the labour market is more readily available based on the European Labour Force Survey, the labour market portrait of European citizens residing in a country other than their own by country of origin or citizenship is less known. Eurostat's new innovative series of products (datasets and explanatory articles) on the characteristics of European citizens residing in another European country by country of origin or citizenship fills this gap. The work is of interest at a methodological level, as the construction of these indicators implies the aggregation of sub-populations coming from different national samples of the European Labour Force Survey. The article discusses the main results (comparing the population living abroad with the one in the country of origin or citizenship but also the one in the destination countries). It also describes methodological matters: the advantages of using this method, the calculation of confidence intervals and data validation using other data sources.

Keywords: emigration; labour market characteristics; EU-LFS; confidence intervals; validation

SESSION 29 – ESTIMATION METHODS TO IMPROVE QUALITY

Title of abstract: Register-based estimation of total dwellings and households

Li-Chun Zhang, Johan Fosen
Statistics Norway

Total of dwellings and households are key statistics of a country. Traditionally these figures are produced based on population and housing census, which can be costly. How to produce these statistics based on statistical data originated from the relevant administrative sources is currently a major challenge to the Census Transformation Programmes at a number of European countries. We study the matter based on the Norwegian Address Register and Population Register. A particular difficulties arises due to the fact that dwelling identification at multi-dwelling addresses is problematic, owing to quality issues in the input sources, despite reliable identification of all the addresses. We develop two extensions to the existing capture-recapture methodology in this context. The first one can be characterised as a two-step approach, where one first applies standard capture-recapture methods to obtain an estimate of the resident addresses (i.e. address at which there exists dwellings and possibly households), and then use various missing-data methods to estimate the number of dwellings and households per address. The second approach can be formulated in terms of a log-linear model, under which it becomes possible to estimate at once the sizes of three populations, namely of resident address, dwelling and resident dwelling (i.e. dwelling household). We demonstrate both the approaches using real-life data. These provide potentially options for purely register-based estimation of total dwellings and households, instead of the costly census.

Keywords: Register-based population size estimation; census transformation programme; capture-recapture of clustered elements; log-linear models

Title of abstract: Applications of multiple tests to improve the quality of multi-source data sets

Czesław Domański, Dariusz Parys
UNIVERSITY OF LODZ, Poland

In official statistics we often use multi-source information like desk researches, administration sources, censuses, reporting and others. It is necessary to estimate these kinds of information. This estimation could be based on making decision of possibilities of integration of these data sets. In our considerations, treating the different data sets as samples, we present the application of multiple testing in estimation of quality of multi-sorce data sets. From statistical experience the stepwise multiple test procedures will be suitable. We use more powerful step-up and step-down procedures which cotrol the error rates in multiple inference.

Keywords: multi-source data sets, multiple testing, stepwise testing procedures

Title of abstract: Exploiting Auxiliary Data: Random forest regression estimator

Luis Sanguiao
National Statistics Institute of Spain (INE)

Suppose we have, in addition to our survey data, some auxiliary data A that are related to our target variables, but we don't know how. The source of the data could be, for example, administrative data or even Big Data, and it covers an important subset of the target population. An important question arises: how to exploit the data to improve our estimations? The increase of the accuracy would allow a reduction in the sample size of the survey. We propose a new regression estimator based on random forests: the Random Forest Regression Estimator. Being similar to the GREG estimator it has the same advantages that random forest regression has over linear regression. In particular, the relationship between A and target variables is learnt directly from data and both discrete and continuous explanatory variables can be used directly. This also means that the method easily accommodates a change in the structure of the auxiliary data. The estimator is nearly unbiased and we give an approximate estimator of its variance for stratified random sampling. It is important to note that the estimator remains unbiased even if the model is poor. Several simulations are run, with both synthetic and real data to show its performance in practice. Finally, several different applications to official statistics are proposed. Aside from sample size reduction, if you use random forests for imputation, you can use the estimator to correct the bias from aggregate means or totals, what constitutes its more interesting secondary use.

Keywords: Multivariate auxiliary information, Sample size reduction, Random Forest, Regression Estimator

Title of abstract: Assessing and adjusting bias deriving from mode effect in mixed mode social surveys

Claudia De Vitiis, Alessio Guandalini, Francesca Inglese, Marco Dionisio Terribili
Italian National Institute of Statistics (ISTAT)

The mixed mode (MM), i.e. the use of different collection techniques in one survey, is a relatively new approach for ISTAT, especially for social surveys. It is adopted both to contrast declining response and coverage rates and to reduce the cost of the surveys. Nevertheless, mixed mode introduces several issues that must be addressed both at the design phase, by defining the best collection instruments to contain the measurement error, and at the estimation phase by assessing and treating the bias effects (mode effect) due to the use of MM, in order to ensure the accuracy of the estimates. Mode effect refers strictly to measurement error differences due to the mode of survey administration, but, when modes are assigned not randomly, a selection effect can generally occur and appropriate inference methods to evaluate mode effect are needed because the two types of error are confounded. Disentangling selection and measurement effects requires auxiliary information that are assumed to be mode insensitive, acquired from registers or collected by the survey itself. The problem of the selection effect can be faced with adjustments based on Propensity Score (PS) approach, allowing to mitigate the confounding effects of the selection mechanism and evaluate correctly the measurement error within homogeneous groups of units. The focus of this work is the experience in the evaluation and treatment of MM effect in the experimental situation of ISTAT "Aspects of daily lifesurvey- 2017", a sequential CAWI/PAPI survey for which a control single mode sample PAPI was planned to make an assessment of mode effect on two independent samples with different

techniques. Methods to assess the impact of MM on the quality of the estimate, the representativeness of the two samples and models to evaluate the measurement error and selection effect in the MM sample are experimented.

Keywords: mixed mode, selection effect, measurement error, propensity score, representativeness index

Title of abstract: Estimation of the standard error for net changes with the EU Labour Force Survey – How can users independently and appropriately calculate standard errors and confidence intervals?

Bernhard Schimpl-Neimanns, Stefan Zins
GESIS – Leibniz Institute fort the Social Sciences, Mannheim, Germany

The EU Labour Force Survey (EU-LFS) is one of the most important official surveys for comparative social research in Europe. As such it is a source for the estimation of indicators to monitoring economic and social policy. To assess if observed changes of indicators are significant or not variance estimation for the estimated changes is required. This task is challenging as most countries use complex sampling designs and different rotation schemes. Due to their partial overlap between waves, rotational panels allow a more efficient estimation of changes. To account for this the covariance of cross-sectional estimates has to be estimated. In practice, users can face some difficulties in doing so because time-consistent identifiers are required. However, the data for scientific purposes released by Eurostat currently contains identifiers for the primary sampling units that are randomized per dataset and are only consistent for one year, supporting the erroneous assumption of statistical independence between waves. By taking the example of LFS-data from Austria we show how the available design information can be used to estimate the variance for change in cross-sectional indicators. For this we use the method proposed by Berger & Priam (2016), which represents a solution to the variance estimation problem in the presence of incomplete sampling design information. Statistics Austria releases files for the Austrian microcensus which can also be used for the Austrian LFS to solve restrictions by longitudinally inconsistent identifiers. This enables an empirical examination of the error, which can occur as a consequence of the erroneous assumption of independent samples. There by we can show how proper variance estimation is feasible. Therefore, we recommend that variables for stratum, clustering, weight, and completely time consistent unit identifiers should be released if there are no confidentiality concerns. This would considerably improve variance estimations based on anonymised microdata.

Keywords: EU Labour Force Survey, Variance Estimation, Linearisation, Net Change

Title of abstract: Adjusting the gender pay gap using the Structure of Earnings Survey data

Piotr Ronkowski, Denis Leythienne
Eurostat

Reducing the gender pay gap (GPG) is one of the key priorities of gender policies at the EU and national levels. At the EU level, the European Commission prioritised „reducing the gender pay, earnings and pension gaps and thus fighting poverty among women" as one of the key areas in its „Strategic engagement for gender equality 2016-2019". The unadjusted GPG, calculated as the relative difference between the average earnings of women versus men, is widely used in this context as the key indicator to monitor and evaluate progress in this area. However, the unadjusted GPG entangles in its measurement both possible discrimination between men and women, in terms of „unequal pay for equal work", as well as the impact of differences in the average characteristics of men and women in the labour market. Against this backdrop, Eurostat has developed a methodology to adjust the GPG using the Structure of Earnings Survey (SES) microdata. The methodology is based on the Blinder-Oaxaca decomposition. The SES microdata provide information on the earnings of individual employees as well as on some personal, job and enterprise characteristics. Eurostat's project provides a decomposition of the difference between male and female earnings into explained and unexplained parts. The explained part is the gap between male and female earnings which is due to the differences in the average characteristics (sector of activity, age, occupation, etc.) of male and female employees. The unexplained part measures the difference between the financial returns to men and women with the same characteristics. Eurostat's methodology and results should help policy makers to better interpret the unadjusted GPG. They should also stimulate further discussion within the European Statistical System on a common method to adjust the GPG indicator.

Keywords: earnings statistics, gender statistics, gender pay gap

SESSION 30 – QUALITY OF STATISTICAL PRODUCTION PROCESSES

Title of abstract: Global standards and national procedures – What can the ESS learn from the Programme for the International Assessment of Adult Competencies (PIAAC)

Markus Bönisch
Statistics Austria

The Programme for the International Assessment of Adult Competencies (PIAAC) is an international OECD survey that compares key competencies of adults (16-65 years) in 33 countries. In order to obtain high quality datathe international PIAAC Consortium produced an elaborate set of standards for almost all aspects of the national implementation. In Austria, a comprehensive set of procedures was put in place for the PIAAC fieldwork. Some of the international requirements for data collection were not reasonable within the national context and required certain adaptations to accomplish a successful fieldwork. The presentation will talk about PIAAC and its methodological background, describe key fieldwork measures in Austria and discuss how specific measures relate to global data collection standards. Reflecting on this national experience, some of the possibilities and limitations of national compliance to global standards will be discussed. Furthermore the multidimensional assessment of quality in PIAAC (Response Rate, Non-Response-Bias, compliance withstandards) will be discussed and related to national contexts. The conclusions will discuss still open issues regarding data quality in cross-national surveys and the balance between global standards for comparability and the degree of freedom to reconcile national differences. Furthermore some best practices and lessons learned will be presented which have the potential to enhance the quality of European statistics and the ESS.

Keywords: global standards, data quality, PIAAC, ESS

Title of abstract: Standardization of Business Statistics Processes in Istat

Mauro Bruno, Giuseppina Ruocco, Orietta Luzi, Monica Scannapieco
Italian National Institute of Statistics (ISTAT)

Since the second half of 2014, Istat has launched a modernization program to improve statistics quality. In order to increase the effectiveness and efficiency of the production chain, the revision and the standardization of the business statistics production processes have become of utmost importance. Currently, business surveys use highly customized methods and tools. Some processes are tied to specific persons, and built on their knowledge and skills. This kind of organization produces two major negative effects: duplicated work and limited reuse of tools and competencies. The Generalized Process for Business Statistics (GPBS) aims to identify and implement a general data model and architecture to standardize similar steps of business surveys. The GPBS initiative has the main objective to standardize:

- methods and tools for all the statistical phases of production processes;
- the workflow, meant as the best combination of data and statistical services to implement the statistical process.

The core of GPBS will be a set of shared and generic corporate services based on structured metadata, for processing, storing and analyzing data. Such services will be designed and implemented enhancing the software currently available in the Istat's generalized software repository. Adopting the conceptual model of the ESS Statistical Production Reference Architecture (SPRA), the main components of the TO-BE integrated architecture, will be services related to Generic Statistical Business Process Model (GSBPM) phases and sub--processes and exploiting enterprise-level building blocks.

Keywords: Process standardization, process integration, statistical service

Title of abstract: Calif – the interactive Shiny web application for calibration of weights of statistical surveys

Boris Frankovič, Katarína Žiaková
Statistical Office of the Slovak Republic

Since many years the Statistical Office of the Slovak Republic has performed calibration of weights in statistical surveys in order to enhance precision requirements and maintain consistency between surveys. Based on our best practice, it has been proved that this procedure has to be carried out with utmost care. To that end, we prepared the first version of Calif with graphical user interface in 2013. This tool enabled users to take advantage of simple environment full of useful statistics that allowed for swift calibration when the most feasible solution was easily found. However, the old-fashioned GUI has found its limits. Therefore, we come with the new interactive web application made in R package Shiny that offers space not only for the enhancement of the calibration process but also for further development. Apart from new graphical user interface that runs in aweb browser, makes calibration an intuitive and convenient process and offers new features, Calif could be used in the future as amore general statistical tool, which would not be limited by just the calibration but also be open for imputation, estimation and variance calculation.

Keywords: calibration, weights, Shiny, R

Title of abstract: Innovation to improve data quality: the case of Italian Household Budget Survey

Achille Lemmi - ASESD - Tuscan Universities Research Centre 'Camilo Dagum', Pisa, Italy
Donatella Grassi, Alessandra Masi, Nicoletta Pannuzi - Italian National Institute of Statistics (ISTAT)
Andrea Regoli - University of Naples Parthenope, Italy

The recent redesign of the Italian Household Budget Survey (HBS) represents the starting point of this contribution whose aim is to assess the impact of both methodological changes and innovation in data collection on the estimates of household consumption expenditures. In 2014, many innovations were introduced in the HBS, in response to changes in European recommendations and purchasing behaviors (in terms of habits and with respect to the choice of goods, services and places of purchase), and to an increased demand for information in the context of social and economic research. New instruments and techniques have been introduced, together with more accurate methodologies, with the aim of improving the estimation quality of household consumption expenditures and of other derived statistics (as the Italian measures of relative and absolute poverty). Because the old and new HBS were conducted in parallel during the last two quarters of 2012 and during all quarters of 2013, it has been possible to compare the effects of the abovementioned

changes on the consumption expenditure estimates. The aim of the paper is not only devoted to the impact of different methodological choices but rather to the implications of differences in data collection and variables definition. Moreover, the impact is evaluated not only in terms of levels, but also taking into account the whole distribution. Also the analysis on which subpopulations are more affected by changes is conducted, by explaining the different effects of each introduced innovation. Finally the paper contains an error profile analysis, with the aim of distinguishing changes due to innovations from possible bias linked to sampling coverage or households' selection. Although our considerations refer to the Italian situation, some evidence and, in particular, the models and analytical procedures may be extended to other realities.

Keywords: data quality, survey methods, consumption expenditure, poverty

Title of abstract: Redesign of the Statistical Information System: Czech experience

Marek Rojicek
Czech Statistical Office

In the last years, preparation and implementation of the integration of software tools for collection, processing and dissemination of statistical data has been realized. The main tasks of this project were decrease of the administrative burden laid down on respondents together with the increased quality and accessibility of statistical data for every user of the statistical information system (SIS). The specific targets of this project, realised within the general framework of e-Government implementation and financed mostly by EU funds, were the integration of mathematical-statistical methodologies, procedures, tools and technologies used for evaluation of users' requirements, preparation and processing of statistical tasks, analysing and publication of statistical information. Within this project the development of data warehouse has been completed including the creation of system of secure and protected access to data, as well as the unification of administrative data sources and introduction of mutual electronic contact with respondents. The main part of the project has been realised in the years 2013 – 2014 and in 2015 the project was transformed into routine statistical production process. After implementation of the new SIS the need for higher cooperation of all statistical as well as IT departments significantly increased. Although most of the roles were anchored into the line organisation structure, it was necessary to keep partly project type of the management of the whole process due to its complexity. We expect that the initial investment in the form of financial resources and human capacities should be returned in the following years by decreased administrative burden of respondents and increased quality and accessibility of statistical indicators for users.

Keywords: metadata, digitalization, integration, data collection

Title of abstract: Specification and immediate visualization of a questionnaire – a metadata-driven approach

Thomas Dubois
INSEE, Paris, France

Insee has been developingduring the last years a metadata-driven questionnaire generator code-named Eno. This tool takes as input a formal description of the questionnaire complying with the DDI standard and executes a completely automated chain of transformations to produce the current survey questionnaire (web and paper questionnaires). The questionnaires can be personalized and, depending on the output format, a variable amount of flow logic can be implemented. Thanks to the automation of the process, Eno refering to DDI standard allows traceability of the changes in the questionnaires and reinforces the soundness of the data collection process. Currently, nearly 15 survey questionnaires were produced by Eno over the last two years, and this figure is expected to grow to more than 40 survey questionnaires by the year 2020. Future work will allow these numbers to grow significantly with Eno providing interviewer questionnaire. An additional module to this generator is a questionnaire design user interface, code-named Pogues, that connects with the generation process. More specifically, Pogues produces the DDI description of the questionnaire which is then submitted to an embedded instance of Eno. Using this tool, a survey manager or questionnaire designer can specify his web questionnaire in a friendly way and visualize the generated result in one click. The current version supports the main functionalities needed for business surveys, and the roadmap foresees the development of more complex logic flows that can for example be found in household surveys. Other future developments will enhance the possibilities to specify questionnaire controls that could be used during the data collection process as well as during the data editing and imputation processes. Pogues is open source and natively internationalized, and in consequence can be used directly by other statistical agencies. Other national statistical institutes have already shown their interest in this software.

Keywords: metadata-driven approach, questionnaire generator, DDI, data collection, innovation

SESSION 31 – STATISTICAL DISCLOSURE CONTROL AND MICRO DATA EXCHANGE

Title of abstract: A new approach for Disclosure control – Random Tabular Adjustment

Steven Thomas, Peter Wright, Mark Stinner
Statistics Canada

The Government of Canada is investing in making more data available to Canadians. Statistics Canada is also investing in this initiative with the way that it assesses and treats disclosure risks. Data has historically been withheld from the Canadian public through a process of cell suppression whenever a disclosure risk has been identified. An alternative risk assessment is being proposed that will rely on ensuring that undesirable statistical inferences are prevented while useful statistical inferences can still be made through tabular perturbation. The Random Tabular Adjustment (RTA) process involves adding random noise to estimates where disclosure risks are apparent. At the same time, we are reassessing what we are comfortable with in terms of risk. The move to releasing more data comes at the expense of taking higher risks. In terms of quality, the new approach entails balancing various aspect of quality, such as pertinence, accessibility and accuracy. This presentation will highlight the challenges with disclosure control, the RTA method will be described and the appetite for risk will be discussed.

Keywords: Disclosure Control, Tabular Adjustment, Confidentiality, Statistics Canada, Perturbation

Title of abstract: Statistical Disclosure Control and Quality Reporting

Arijana Amina Ramic
Wiesbaden, Germany

The role of quality reporting is to demonstrate that high quality standards have been applied and achieved through in the statistical production processes, and to guarantee correct interpretation and use of the produced statistics. Statistical Disclosure Control (SDC) refers to the measures taken to protect data in accordance with confidentiality requirements, ensuring at same time that the usefulness of the data outputs is preserved to the greatest extent possible. Disclosure control measures reduce data quality (i.e. by suppressing data or changing detail levels), can affect the accuracy of information released (i.e. by data perturbation) or produce bias (i.e. using value rounding or noise addition), and limit access to certain groups (such as researchers). The degree and method of disclosure control may vary for different types of outputs as well as different statistics producers. The peer reviews in the European Statistical System (ESS) show that the highest standards for protection of the statistical confidentiality are applied across the ESS. Some quality indicators, as specific and measurable elements of statistical practice used to characterise the quality of statistics when possible, are also proposed for the information about the statistical confidentiality and security. Nevertheless, for the time being there is no information in the quality reports in the ESS about any disclosure measures applied or their influences at the quality of statistics produced and published. This paper focuses on showing the importance of the presentation of the SDC methods applied and their impacts at the statistics produced, in the quality reports. It provides a small analysis of the bene-

fits for different kinds of the statistics users and of some possible and relevant quality indicators informing about SDC. Some information on the SDC measures applied and/or planned for the Population Censuses 2011 and 2021 will serve as example of good practices of reporting about SDC.

Keywords: Quality Reports, Statistical Disclosure Control (SDC), Quality Indicators on SDC

Title of abstract: Statistical confidentiality: New initiatives in the European Statistical System

Aleksandra Bujnowska, Wim Kloek, Fabian Bach
Eurostat

The protection of confidential information has a huge impact on how statistical data can be published and used for analysis, which makes it a key aspect of data quality. This paper presents new methods and tools currently being investigated in the ESS in order to publish more – and more useful – data without compromising statistical confidentiality. It covers new methodological and IT developments, where concrete use cases demonstrate their impact on data quality. For instance, a promising methodological direction is random noise: several ESS use cases at different maturity stages are presented, including recommendations for the harmonised protection of 2021 EU Census data. We also show how all these developments will enhance the user experience at various levels.

Keywords: statistical disclosure control, cell key method, adding random noise

Title of abstract: Data exchange between the Nordic countries – supplementing education registers by qualifications completed in another Nordic country

Kaija Ruotsalainen, Nicola Brun
Statistics Finland

High-quality and precise statistics production concerning the educational structure of the population usually requires a comprehensive register of qualifications that includes data on all degrees and qualifications completed by the population. An education register is the basis for producing education statistics and also important source data, for example, for population census and various survey-based statistics. Both Finland and other Nordic countries have a statistical register of education that contains data on degrees and qualifications completed by the population. The quality of the education registers is high as a rule because educational institutions are obliged to deliver data to the education registers annually. A big problem is caused, however, by the lack of data concerning qualifications attained abroad. No comprehensive register-based data source exists with data on the qualifications of immigrants attained in their home country nor on qualifications attained abroad by the original population. As part of the project "Nordic Mobility" funded by the Nordic Council of Ministers, the coverage of education registers is improved by exchanging unit-level data between Nordic statistical institutes. This is enabled by the EU Statistics Act, based on which confidential data can be released from one ESS authority to another in order to develop and produce European statistics and improve their quality. This type of data exchange is likely to be the first of its kind at least in the history of Nordic statistical institutes. It is assumed that the data exchange will improve the coverage of the education registers of Nordic countries.

At least in Finland this is likely to be visible especially in Åland where many of the population have completed their qualifications in Sweden, which means that data on completed qualifications degrees have not been included in Statistics Finland's education register.

Keywords: Education register, micro data exchange

Title of abstract: Quality assurance in micro-data exchange – The international trade in goods statistics as concrete example

Anne Berthomieu
Eurostat

The objective pursued when exchanging micro-data is to benefit from an additional data source with no extra burden for the providers of the statistical information (e.g. enterprises). The backbone of a micro-data exchange is a performant, robust and secure transmission system, embedding comprehensive guidelines and binding rules for both the sender and the receiver of the information. Rules binding the receiver are mainly targeting the respect of the data confidentiality while the sender is mainly bound to quality-related targets covering the data completeness, accuracy and punctuality. The quality monitoring is based either on information directly derived from the exchanged datasets or on related metadata. The quality assurance is built on five pillars:

- a requirement system composed of binding rules on who should do what, how and when;
- a guidance system aiming at promoting good practices and improving data harmonisation;
- a reporting system composed of all types of metadata to be attached to the data flow or to be provided as an additional component;
- a monitoring system gathering all information (e.g. error reports) on the data flows and the quality of the exchanged data; and
- an assessment system producing compliance reports pointing out where actions are needed.

Such quality framework is on the way to be implemented for the exchange of micro-data relating to intra-EU exports of goods. It takes its inspiration from the framework already in place to ensure the quality of the monthly trade in goods statistics issued from those micro-data and disseminated by Eurostat. The new aspects to be taken into account are inherent to the high sensitivity of micro versus macro data as well as to the challenge to make the exchanged data as useful as possible for the receiving partner countries.

Keywords: quality assurance, quality requirements, quality monitoring, quality reporting, quality assessment

SPECIAL SESSION 32 – ADDRESSING OS?JGRW IN THE CONTEXT OF BIG DATA – THE ESSNET CONTRIBUTION

Title of abstract: Quality evaluation of experimental statistics produced by making use of Big Data

Giulio Barcaroli, Natalia Golini, Paolo Righi
Italian National Institute of Statistics (ISTAT)

In 2017 the Italian Institute of Statistics (Istat) has started the production of a set of experimental statistics based on the use of Internet data, one of the most relevant Big Data sources. These statistics refer to the activities that enterprises carry out in their websites (web ordering, job vacancies, link to social media, etc.) and are a strict subset of those currently produced by the "Survey on ICT in enterprises". The idea is to calculate these estimates by making use of the websites content, that is collected by using web scraping tools, and processed by applying text mining techniques. Then, models are fitted in the subset of enterprises for which both survey reported values and relevant terms obtained by the web scraping/text mining procedures are available. Experimental statistics have been obtained by making use of two different estimators: the first one is a full model based estimator; the second one is an estimator that combines model based estimates and survey estimates. Considering the various domains for which they have been calculated, the three sets of estimates (survey, model and combined) in most cases are not distant (i.e. model and combined estimate values lay in the confidence intervals of survey estimates), but in some other they do are.

The question is: how to evaluate the accuracy of the three sets in order to understand if experimental statistics can substitute survey ones?

Considering the different factors that can produce bias in survey estimates (total non-response and response errors) and in alternative estimates (population under-coverage and prediction errors), these factors are analysed in detail with respect to the real conditions in the 2017 experience. Finally, a simulation study is carried out in order to investigate the conditions under which a given estimator performs better than the others.

Keywords: Big Data, Internet data, official statistics, model based estimation, quality evaluation

Title of abstract: AIS: Defining a ship's journey and sea traffic analysis

Christina Pierrakou - Hellenic Statistical Authority (ELSTAT)
Tessa de Wit, Marco Puts, Anke Consten - Statistics Netherlands

Ships broadcast information on their location and status on a frequent basis by means of a radio signal. This so-called Automatic Identification Signal (AIS) provides a big data source for maritime and emission statistics. Research on the use of AIS for official statistics is part of the ESSnet Big Data project, performed by The Netherlands (Work package leader), Denmark, Greece, Norway and Poland. Here, we present methods to define a ship's journey and to perform sea traffic analyses. Defining the journey of a ship by using AIS data is needed to obtain insight into all the ports the ship visited. It is also makes it possible to calculate the distance a ship travels. Furthermore, a ship's journey is necessary to improve the calculation of emissions. To determine the journey of a ship, we further build on the algorithm we already developed to determine

a port visit. The resulting journey algorithm is output-driven and enables us to define the start of a journey and to deal with noise in the signal. For traffic and economic analyses, we also wanted to explore the possibility of calculating the number of ships during a certain time interval at certain coordinates by using AIS data. We calculated the traffic intensities of ships around Europe. An important prerequisite is that the grid elements all should have the same size, thus the grid was defined as areas of 10,000 square kilometres. Different coordination systems were compared, with the WGS1985 system being selected. To draw the data to a map, the Lambert Azimuthal projection was used. This method preserves surface area under the transformation. The final visualization is done in Shiny in combination with Leaflet.

Keywords: AIS, Big data, Maritime statistics

Title of abstract: Big Data quality issues regarding multi-domain statistical data combining – a survey and case studies

Jacek Maślankowski - University of Gdansk/Statistical Office in Gdansk
Anna Nowicka, Janusz Dygaszewicz - Statistics Poland

Employing Big Data methods and tools to produce statistical data makes the necessity of the revision of the data quality framework for official statistics. Although many different efforts have been made, including UNECE Big Data Quality Framework or different approaches in research papers, there is no unified Big Data quality framework that can be applied for different type of data sets, such as social media or large structured data sets. On the other hand, the variety of Big Data quality frameworks allows creating the set of quality indicators that will assess different aspects of the data source usability. Therefore, the solution is to create different frameworks depending on the data set used. It is rather easy when one dataset is used. More complicated is when different data sets are integrated, including various data types. The aim of the paper is to show how Big Data quality frameworks can be applied to create the set of indicators that will allow assessing the data set quality in three different stages – as input data sources, during processing phase (data sources integration) and when producing the output data (final experimental tables). The paper covers different aspects of Big Data integration. The first is intra-domain when combining data sets within three different statistical domains: population, tourism and agriculture. It includes combining data from traditional surveys and Big Data sources, such as social media data or satellite data. The second aspect is when combining inter-domain data sources. We have tested data integration by combining population and tourism data sets. The case studies and pilot surveys allows creating original conclusions on how to measure the data quality and which quality indicators can be applied to provide reliable assessment of the data sources and results.

Keywords: big data, quality framework, data integration, data combining

Title of abstract: Coverage of AIS data: comparison of privately held to national datasets – Poland and Hellenic experiences

Christina Pierrakou, Eleni Bisioti - Hellenic Statistical Authority (ELSTAT), Michał Bis - Statistics Poland, Anna Bilska - Statistical Office in Szczecin

AIS data, real-time measurement data of ship positions, is one of the potential Big Data sources investigated by the ESSnet Big Data project. The main aim of the specific work package (WP4) is to explore the potential use of AIS data in the production of official statistics, due to their advantage of being generic worldwide and obtainable at European level. Five National Statistical Institutes participate in WP4: the national statistical institutes of the Netherlands (Work package leader), Denmark, Greece, Norway and Poland. This paper outlines the results of the comparison of national public AIS data to privately held AIS data, in terms of quality and metadata. Various Big Data technologies were used, to store, manage and process the huge volumes of AIS data such as the Distributed File System of Hadoop, the large-scale data processing engine Apache Spark, Scala language, the noSQL database Elasticsearch using GeoData. Moreover, the exploration and visualization tool Kibana was used. The main conclusion drawn from this paper is that AIS is a big data source with great potential to improve official statistics, however more work is needed in the area of ensuring transparency and soundness of methods and processes for the privately held data to be incorporated in the statistical production.

Keywords: AIS, position data, quality of big data source, big data technologies

Title of abstract: Estimation of population counts combining official data and aggregated mobile phone data

Ciprian Alexxandru, Bogdan Oancea, Benjamin Sakarovitch - Romanian National Institute of Statistics (INSSE)
Marc Debusschere - Statistics Belgium
Elisa Esteban, Soledad Saldaña, David Salgado - National Statistics Institute of Spain (INE)
Ossi Nurmi, Pasi Piela - Statistics Finland
Roberta Radini - Italian National Institute of Statistics (ISTAT)
Luis Sanguiao - Statistics Madrid
Martin Tennekes - Statistics Netherlands (CBS)
Susan Williams - Office for National Statistics (ONS), London
Markus Zwick - Statistisches Bundesamt (DESTATIS)

Beyond doubt, mobile phone data stand as one of the most promising Big Data sources for the production of official statistics. In consonance, in the recent ESSnet on Big Data participated by 22 partners of the European Statistical System (ESS) a work package was completely devoted to the access to these data, the development of statistical methodology, the analysis of IT tools and of quality issues to make this promising information source become a regular resource in the production of official statistics. We offer a summary of the works conducted in this work package, going from the intricate issue of accessing diverse forms of mobile phone data (microdata/aggregated data) over setting up an inferential framework to use aggregated mobile phone data in combination with official data to produce population counts, to the development of some IT tools for providing a proof of concept and first analytical results upon real data. All these enter as relevant factors in the quality assessment of the final estimates. As explained in the results of the ESSnet, although we have been able to collect enough real data as to conduct the analytical study, the access to mobile phone data is still an open question which needs further work within the ESS and the European Union. A first set of conclusions and guidelines for partners of the ESS have been obtained. Regarding the statistical methodology, unable to use traditional survey sampling techniques, we have explored the use of hierarchical statistical models as in ecological sampling to propose a generic inferential framework for the counts of diverse target populations (commuters, resident tourists, inbound tourists, general population,. . .). The analysis is completed providing software tools to implement this methodological proposal, showing a proof of concept with both simulated and real data, and assessing the quality of the final estimates.

Keywords: mobile phone data, access, hierarchical modelling, software tools, quality

SPECIAL SESSION 33 – COMMUNICATION QUALITY AND ENGAGING WITH USERS

Title of abstract: Branding of official statistics: new evidence and recommendations

Martina Hahn, Christine Kormann, Kirsi Karkkainen
Eurostat

Better communicating on the value of official statistics has been identified as a strategic need for statistical organisations. It is one of the key goals of the ESS Vision 2020 and it is implemented through the DIGICOM project (Digital Communication, User Analytics and Innovative products). In this context, Eurostat has commissioned a study to get a better understanding of users' perception of official statistics, as well as strategic and operational recommendations on communication. The study will look at general branding aspects, such as brand awareness (do users know the various brands) and brand positioning (how do users perceive the brand as regards its competitors?). It will also cover users' views on the quality of European statistics (Do users associate European statistics with high quality? Which quality aspects are important for them? Are there quality gaps?). This study will be carried out from December 2017 to June 2018 in 8 EU Member States. It will mainly rely on qualitative methods (focus groups and in-depth interviews of representatives of the main user groups), but will also include e-reputation analysis and a quantitative survey. This paper will present the methodology and early results of this study for Eurostat and the ESS as a whole.

Keywords: Communication, branding, user perception, value of official statistics

Title of abstract: Engaging with users to modernise the dissemination of European statistics

Maja Islam, Julia Urhausen
Eurostat

Modernising the dissemination of European statistics is driven by users' needs and facilitating access to official statistics. Following current trends, more visual and attractive content is used, while text is presented in a structured and concise way, addressing the most common user questions. Engaging with users serves as the fundament and impetus for any changes in this process. In 2017, several user research activities were launched at Eurostat as part of the DIGICOM project – an ESS project aiming to modernise the dissemination and communication of European statistics. The aim of these user research activities was to learn more about our users and their needs, and get recommendations on what we can do to modernise the dissemination of European statistics. Two methods were used: field studies and usability tests. In the field studies, 40 different users were interviewed over a period of 6 months. Users were asked about their profile and their use of statistics, and observed as they interacted with a number of dissemination products on the Eurostat website. The outcomes were high-level recommendations on how to improve the dissemination products tested, and personas of the users of European statistics. The usability tests, conducted with smaller groups of users and focusing on a smaller number of products, resulted in more specific recommendations to improve the usability of the tested products. In practice, this is a circular process: Eurostat proposes new or improved dissemination products to users who then provide their feedback; on this basis, recommendations are made

which subsequently result in improvements of the products. Learning from users now will help Eurostat to disseminate better, custom-made products in the future. This presentation will include concrete examples of user feedback and its translation into improved dissemination services.

Keywords: user engagement, dissemination, European statistics, website, user research

Title of abstract: What do ONS users want from quality and methods information?

Sarah Tucker, Alison Davies,
Office for National Statistics, Newport

"There's no substitute for talking to real users"

As National Statistical Institutes (NSIs), we must provide statistical quality and methods information that helps our users make better use of our data. We decided to review our standard user-orientated quality report (Quality and Methodology Information (QMI)) to ensure that it continued to meet user needs, the first question we asked was "How can we find out what our users really need from our quality and methods information?" The simplest answer, though not necessarily the simplest thing to do, was to ask them. With the help of an ONS user researcher we ran a series of user tests designed to:

- Discover who our main users are
- Help us understand which of the current QMI topics are most important to users
- Help create Quality Information User Profiles

Gain detailed insight on what our users need from quality and methods information and what they use this information for. Discover if additional topics are needed to help users make better decisions about the data. Test if our interpretations were correct when redesigning the contents of the QMI. In this paper I will describe the methods we used to meet these goals and discuss the main findings from the user tests. I'll share our newly created Quality Information Profiles which give insight on what different types of users need from quality and methods information, to enable them to use the data with more confidence. I'll give some detail about our users and what they told us about their needs. This included a few surprises which required us to do some rethinking of our assumptions to ensure that we stayed on the right path. Finally, I will share the new contents list and briefly discuss the next steps.

Keywords: Quality, users, communication, engagement

Title of abstract: Users' engagement and national quality reporting at Istat

Andrea Bruni, Maria Francesca Loporcaro, Giorgia Simeoni
Italian National Institute of Statistics (ISTAT)

Istat has recently re-organised the set of activities carried out to manage the relationship with users. The resulting framework is aimed at segmenting users (identification of profiles in order to classify users with different needs), collecting information needs, assessing the overall users' satisfaction. A wide set of methods and tools are used to fulfil these objectives, ranging from traditional satisfaction surveys, to consultations on specifics topics, or indirect analysis of users' requests and feedbacks. The paper will firstly present this new comprehensive approach as well as the tools already in use. Secondly, the focus will be shifted to the results of users' consultation in terms of perceived quality of official statistics and metadata communication. Indeed, since 2014 the yearly Istat users' satisfaction survey includes also two sections devoted respectively to disseminated metadata and product quality. Then, recent developments on the tools for reporting quality to users, that would also meet the users' needs according to findings of previous analysis, will be presented. In particular, the new detailed national Quality Reports, harmonised with ESS metadata standards and oriented to expert and demanding national users, are currently being implemented. They fill a gap in the reporting for the users, that so far has been oriented to not expert ones, through the dissemination of Quality at a glance reports. They have also been developed in response to the recommendation of the last round of Peer Review on the ES Code of Practice. So far, the Istat quality reporting tools are tailored for traditional statistics based on survey and/or administrative data, however the issue of documenting the new experimental statistics is arising. A preliminary hypothesis on how to document this kind of statistics will be finally proposed, also on the basis of a review of what is already done in other statistical institutes.

Keywords: quality reports, perceived statistics quality, dialogue with users, communicating quality

SPECIAL SESSION 34 – IMPROVING THE QUALITY OF MULTI-MODE DATA COLLECTION. THE EUROPEAN MIMOD PROJECT

Title of abstract: The MIMOD Project: an overview

Marina Signore
Italian National Institute of Statistics (ISTAT)

In response to a Eurostat call for a Grant on "Cooperation on Multi-Mode Data Collection", a Consortium led by Istat (Italy) in partnership with CBS (Netherlands), SSB (Norway), STAT (Austria) and Destatis (Germany) was set up and the MIMOD – Mixed Mode Designs for Social Surveys was awarded the Grant. A network of supporting countries - INSEE (France), Czech Statistical Office (Czech Republic), Central Statistical Office of Poland (Poland), Statistic Finland (Finland) and Statistics Sweden (Sweden) - will provide inputs to the project. The MIMOD project aims at supporting NSIs in facing a range of challenges which are at the forefront of applied research when implementing multi-mode and multi-devices data collection. The MIMOD project covers the following main topics: I) mode organisation with the objective of determining the steps in a decision tree that supports the implementation of adaptive and responsive mixed-mode survey designs; II) mode bias/mode effect and its adjustment with the aim of providing general guidelines on methodologies to deal with mode effects in multi-mode designs; III) case management in mixed-mode data collection with the purpose of investigating the different systems in use in terms of technical components and organisational approaches used, as well as challenges in efficiency and quality ; IV) mixed-mode questionnaire designs in order to give best practice recommendations on approaches for developing questionnaires for mixed-mode surveys as well as on modes used in the contact and follow-up phases of data collection; and V) challenges for phone and tablets respondents in CAWI with the aim of investigating the use of mobile devices (smartphones, tablets) in ESS surveys, and of mobile device sensors (such as GPS, camera, microphone, accelerometers) to enrich ESS surveys.

The paper will describe the activities and the mid-term results achieved by the MIMOD project.

Keywords: multi-mode data collection, mode effects, mode organisation, case management systems, multi-devices data collection

Title of abstract: Dealing with mode effects

Orietta Luzi - Italian National Institute of Statistics (ISTAT)
Bart Buelens - Statistics Netherlands

One specific workpackage of the European project MIMOD - MIxed MOde Designs for social surveys is devoted to the evaluation of methodologies for dealing with mode effects in surveys using multi-mode data collection. Recent literature widely discuss the problem of identifying and treating mode effect components, i.e. mode selection (resulting from errors of nonobservation), and mode measurement (resulting from observation errors - coverage, nonresponse, and measurement errors). The purpose of the workpackage is to look into ways to cope with mode effects (e.g. weighting, imputation, other methods), and to analyze differences in the final sample composition based on different modes across time, countries and survey types, providing practical, evidence based guidelines for the NSI's. In the first period, the activities will be focused on

an updated literature review on methods for the assessment and the adjustment of mode effects in mixed--mode designs, particularly those currently used in the ESS, with a discussion of assumptions, advantages and disadvantages of the various approaches. This activity will be supported by the survey which will be conducted at the beginning of the projcet on a selected group of NSI's. The suitability of selected statistical approaches and methods for mode effect assessment and adjustment will be then evaluated based on applications on current multi-mode social surveys. Based on the previous activities, general guidelines about methods and approaches which can be adopted to deal with mode-effects in multi-mode designs will be provided. The paper will report some outcomes of the workpackage's activities in the first half of 2018, including some results of the MIMOD query and main evidences from literature review. The overall strategy which will be adopted in order to assess mode effect estimation (and adjustment) in selected ESS surveys will be also delineated.

Keywords: multi-mode data collection, mode effects

Title of abstract: Adapting ESS survey questionnaires to mixed-mode data collection

Dag F Gravem
Statistics Norway

Many countries in the European Statistical System (ESS) are or have recently been in the process of transferring previous interview and paper surveys into web and mixed-mode surveys. A 2012 survey among 17 National statistical institutes (NSIs) showed that this process for many NSIs involved major adjustments of both wording, structure and placement of instructions. Shifting from interviewer administered surveys to self-administered in particular potentially can result in mode effects resulting in differences in estimates. No structured overview of such question changes or the effect they may have on data quality for key questions and surveys within the ESS exists. Work package 4 of the MIMOD (Mixed Mode Designs for Social Surveys) project aims to amend this by offering evidence-based recommendations for mixed-mode questionnaire adaption. This will be done by reviewing available documentation from NSIs, as well as through prestesting conducted specifically for the project. The paper will present the work package in more detail, including which surveys, questions and question types that are analysed and tested, and offer preliminary results from pretesting.

Keywords: Mode effects; Social surveys; Data collection

Title of abstract: ESS surveys and mobile device data collection

Barry Schouten - Statistics Netherlands
Bart Bakker - Statistics Netherlands and VU University Amsterdam

Mobile device coverage has increased steadily and to data a substantial share of respondents attempts to complete online surveys on smartphones and tablets. Surveys in the European Statistical System have, however, rarely been made fit, let alone been optimized, for such devices. Apart from the completion of surveys on mobile devices, the devices also have another feature: They carry various sensors that facilitate automated measurements once respondents consent. In a number of surveys such as travel and health, such sensor measurements have already been tested and applied. In the paper, we discuss the application of mobile devices to ESS surveys: Are the surveys fit to be completed on mobile devices? Or can they be made fit with reasonable effort? And are sensor measurements useful and added value for the ESS surveys?

Keywords: Mobile device; Survey; Sensordata

SPECIAL SESSION 35 – INSTITUTIONAL FRAMEWORK AND IMPLEMENTATION OF QUALITY MANAGEMENT SYSTEMS IN THE ENLARGEMENT AND THE EUROPEAN NEIGHBOURHOOD POLICY (ENP) COUNTRIES

Title of abstract: Key elements of a Quality Management System:Example of Georgia (institutional set-up of a quality framework, quality culture)

Levan Tserediani, Teimuraz Gogishvili
National Statistics Office of Georgia

The Adapted global assessment mission conducted jointly by Eurostat and UN ECE Statistical Division in 2012 recommended an introduction of quality management system at Geostat and the use of the ESS quality parameters for monitoring the quality of statistical outputs. In May 2014 a Methodology and Quality Management Unit (MQMU) was created at Geostat. The main area for this unit is to handle quality issues within all areas of Geostat activities. It also includes quality culture, naturally emphasizing continuous improvement of processes and results in satisfied customers and improved image of Geostat. Creating a full-time unit dedicated to quality issues was a key institutional element for the systematic monitoring of quality. Based on the desk research on international standards the EU model as a priority for Quality Assurance Framework based on the ESS Code of Practice (CoP) has been selected. The quality component within the cooperation project with Statistics Sweden gave Geostat a good foundation for future work in that area. MQMU has taken some active steps in monitoring and improving quality. Quality mapping was conducted for all processes, resulting in prepared reports and recommendations; standard routine descriptions based on GSBPM has been prepared; quality reports (based on ESMS) writing process has started; several policy documents and guidelines on statistical data confidentiality and revision have been prepared. In order to facilitate the quality culture spread through the whole statistical system an interagency working group on quality issues has been created. Training courses on CoP and quality assurance issues for other producers of statistics and data providers have been conducted. The paper will talk about the challenges and problems experienced in setting-up the MQMU and in spreading a quality culture in Geostat as well as some results of this work.

Keywords: Quality assurance framework, quality culture

Title of abstract: Implementation of a quality management system in the NSI of Serbia – success stories and future plans

Nataša Cvetković, Miodrag Cerovina
Statistical Office of the Republic of Serbia

During the last several years, and especially after the Light peer review exercise (2011), Statistical Office of the Republic of Serbia (SORS) made significant efforts in order to introduce quality management system. SORS is highly devoted to quality, and this could be seen from various conducted activities. Information and knowledge about different quality issues are continuously being spread among all employees through intranet portal on quality and different educational activities. User and staff satisfaction surveys have already being implemented for certain number of years. Documentation system for quality management, based on ISO 9001 standard, has been introduced, as well as a system for producing reference metadata and quality reports (RZSMETA). Standardization of the production process started in SORS many years ago, initialized by introduction of Integrated Survey Technology (IST). Completely designed and developed in SORS, IST has become regionally recognized data integration concept that supports several phases of statistical production process (build, collect, process and analyze). It is a challenge now how to efficiently use this comparative advantage in order to support building general metadata system as a precondition for further improvement of statistical process and product quality. The paper will describe how these different elements of a quality management system were implemented and what problems/solutions were identified/found by the Serbian statistical office.

Keywords: quality management, metadata, standardization, IST

Title of abstract: Quality management in the statistical office with a view to implement the CoP

Sigalit Mazeh, Moshe Pollak, Batia Attali, Daniel Roash
Israel Central Bureau of Statistics

Quality of statistics is a catch-all phrase for a number of concepts, entailing production, standardization, documentation and publication of statistics. The Israel Central Bureau of Statistics (ICBS) is in the process of trying to advance quality in all of these fronts. First and foremost, quality in production involves persuasion of universal implementation of a Code of Practice. Standardization requires uniformity of definitions and documentation across offices that disseminate statistics. Although every producer of official statistics in Israel is required by law to consult with the National Statistician, in practice this has heretofore not been religiously applied. This talk will present past efforts and present state of affairs regarding the endeavor of the ICBS to streamline statistics in Israel and bring it up to modern quality standards. The talk will describe the efforts expended to obtain the cooperation of disparate governmental agencies as well as internal efforts to modernize procedures in the ICBS.

Keywords: Quality management, Code of Practice, National Statistical System

Title of abstract: Challenges of applying the model of total quality management in the Statistical Office of Montenegro

Gordana Radojevic, Jelena Markovic
Statistical Office of Montenegro

Recognizing the experience of Statistical Office of Montenegro in harmonizing national statistical systems with the standards of European statistical system, which are related to the institutional and methodological harmonization, the hypothesis of this paper is that for the successful selection and implementation of a model of quality, of great importance is first to understand each aspect of statistical process, and after that the implementation of model that will make this statistical process more qualitative. Therefore, the aim of this paper is to analyze the link between the models of education of the employees in the Statistical Office in order to understand the statistical processes in context of EU integrations, and then explain why and how the TQM model is implemented in the Statistical Office of Montenegro and thus conclude with the key challenges of Montenegro as an enlargement country in the implementation of TQM.

Keywords: Total Quality Management model, education, statistical process

Title of abstract: Key elements of a quality management system in the statistical office of Ukraine

Vadym Pishcheiko, Anna Rozumna
State Statistics Service of Ukraine

In 2010 the State Statistical Service of Ukraine (SSSU) approved National principles of the activities of state statistics bodies, harmonized with the European Statistics Code of Practice and adopted the Generic statistical business process model (GSBPM) as a platform for describing the process of statistical production. At the initial stage, descriptions of all statistical surveys were standardized in a unified form. The process approach to monitoring and evaluating performance has been introduced since 2012. An annual inventory of statistical surveys was conducted identifying problems in detail at the level of the survey, defining solution and future tasks for all statistical surveys. During the inventory, proposals and ideas from functional and territorial bodies, suggestions from respondents expressed during the filling in of the questionnaires as well as from users are taken into account. To monitor activities, the SSSU uses four self-assessment questionnaires:

1. Self-assessment questionnaire based on the DESAP (Development of a Self-Assessment Program);
2. Self-assessment questionnaire on improving statistical supervision (for the heads of statistical supervision);
3. Self-assessment questionnaire on the results of the activities of the territorial bodies;
4. Self-assessment questionnaire on the functioning of the processes of statistical production (for authorized employees, responsible for implementing procedures and operations).

In 2013-2015, the SSSU started introducing a process approach into planning. A statistical survey programme is prepared based on the process model, which allows standardizing the description of procedures and operations performed within statistical surveys. In order to unify approaches to the preparation of quality reports, the SSSU in 2014 developed Recommendations on the preparation of a standard report on the quality of statistical surveillance, which complies with the ESS Quality Report 2009, and now all statistical surveys are covered by standard quality reports.

Keywords: quality reporting, self-assesment, comprehensive approach to quality, business process model

SPECIAL SESSION 36 – QUALITY ASSURANCE OF STATISTICS UNDERLYING THE MACROECONOMIC IMBALANCE PROCEDURE (MIP)

Title of abstract: The importance of high quality statistics for the MIP

Alessandro Turrini, Vitor Martins
DG ECFIN, Luxembourg

The Macroeconomic Imbalance Procedure (MIP) is a system for monitoring macro-economic developments and policies and detecting potential harm to the proper functioning of the European Union (EU). It is part of a surveillance system for budgetary and economic policies, implemented via the European Semester of economic policy coordination, the EU's economic policy-making calendar. An essential tool in the procedure is the MIP scoreboard, together with a set of auxiliary indicators. The scoreboard is designed to capture the most important internal and external aspects of macroeconomic imbalances through a limited set of relevant indicators of high statistical quality, ensuring the maximum possible degree of consistency and comparability among statistical domains and across countries. The MIP scoreboard has changed over time, due mainly to the availability of better and more comparable data, and to the inclusion of some social indicators. Efforts made by statistical producers, in the European Statistical System and in the European System of Central Banks, to improve the quality of MIP indicators are highly valuable.

Keywords: Macroeconomic Imbalances Procedure

Title of abstract: Ensuring quality of Macroeconomic Imbalance Procedure (MIP) indicators: a multi-level cooperative approach

Silke Stapel-Weber - Eurostat
Caroline Willeke - European Central Bank, Frankfurt am Main, Germany

The credibility of the MIP and its smooth implementation depends on the availability and quality of the statistics underlying the MIP. Statistics used in the procedure have to be fit for purpose, reliable, timely and comparable among Member States. For that reason, the quality of MIP indicators is constantly monitored and improved. Statistics underlying the MIP indicators are based on European Union legislation and compiled by the European Statistical System (ESS) and the European System of Central Banks (ESCB). In order to ensure the best possible quality, Eurostat and the Directorate General Statistics of the European Central Bank have jointly developed a quality assurance framework based on a three-level structure, covering country specific reports, in the form of self-assessments, domain-specific quality reports, and an annual ESS-ESCB quality assessment report on MIP statistics which is addressed to the European Parliament and Council, policy makers and the public at large. The report assesses the reliability and comparability of statistics underlying the MIP and identifying measures to tackle most relevant quality issues. A Memorandum of Understanding signed by the two institutions in November 2016 has established practical, cooperative approaches for statistics from Balance of Payments and Financial Accounts. This paper will explain the background to ESS/ESCB cooperation on the quality assurance of statistics underlying the MIP, and describe the steps being undertaken to implement this assurance.

Keywords: Macroeconomic Imbalances Procedure

Title of abstract: Reconciliation International Investment Position and Financial Accounts in Belgium

Rudi Acx, Roger De Boeck
National Bank of Belgium

In Belgium, the National Bank of Belgium is responsible for compiling both financial accounts and balance of payments statistics, and therefore has a key role in providing statistics for MIP purposes. The bank has worked for many years to continuously improve the quality of its statistics and has actively participated in European-level quality initiatives. The paper will provide the background to the work in Belgium and some reflections from a national level on recent developments in the European framework.

Keywords: Macroeconomic Imbalances Procedure

SPECIAL SESSION 37 – QUALITY OF MULTI-SOURCE DATA

Title of abstract: Quality Guidelines for Multisource Statistics

Giovanna Brancato, Antonia Boggia, Gabriele Ascari
Italian National Institute of Statistics (ISTAT)

The Guidelines on the quality of multisource statistics are being produced in the framework of the ESS.VIP Admin project, as a work package of the Essnet Quality of multisource statistics. They are intended to be practical and applicable in the National Statistical Institutes of the ESS, when several administrative sources are used exclusively or in combination with survey data to produce statistics. The guidelines are addressed to process managers who can use them in the design and the implementation of multisource statistical processes, to identify the factors having higher impact on output quality and to monitor and assess quality. First of all a review of existing quality guidelines published at national level has been carried out, with the scope of identifying the elements relevant for the development of guidelines within the project. The structure and the content of the guidelines have been defined and reviewed taking into account the comments from Eurostat and the ADMIN network of contact points. A draft version of the guidelines is expected by the end of the current Specific Grant Agreement, i.e. in September 2018, however some areas, for which some methodological work is still required, are expected to be further finalised in the framework of the third specific grant agreement of the Essnet. The paper will describe the quality framework adopted in the guidelines, will present the structure of the guidelines and a sample of the content. The guidelines are structured according to Eurostat statistics quality dimensions and take into account: the different basic data configurations when using multisource data, the phases and sub-processes where the errors generate, examples of measures and quality indicators. This work is carried out in collaboration among the members of the Essnet.

Keywords: output quality, statistical process error sources, multisource statistics

Title of abstract: Administrative data and quality, guidelines towards better quality of administrative data

Sorina Vâju, Mátyás Mészáros
Eurostat

Statistical authorities need to produce data faster in a cost effective way, to become more responsive to users´ demands, while at the same time providing high quality output. One way to fulfil this is to make more use of already available data sources, and in particular administrative sources, most typically used in combination with other sources. Depending on the use of the administrative sources and the data configuration different statistical tasks must be applied. Usually it is not only one task but a sequence of different tasks that have to be applied, for example, data integration, imputation and editing or tabulation. For these tasks different methods are available and depending on the input data quality and the data configuration the same method can have limited use or produce lower quality outputs. The use of administrative data sources risks impacting negatively quality on several dimensions, in particular accuracy and comparability. Surveys and administrative sources have both particular strengths and weaknesses. Combining them may overcome these weaknesses, provided that suitable methodology and tools are used. At the same time, harmonised measures of quality for outputs that combine administrative sources with other sources (surveys) are neces-

sary to ensure that European Union official statistics are of sufficient quality and fit for their intended use. This paper looks at the most frequent methodological challenges faced when integrating administrative sources and provides, for typical situations, preferred methods to have the best quality of statistical output. It also introduces the work of ESSnet on the Quality of Multisource Statistics (KOMUSO) to develop quality measures and guidelines related to the use of administrative sources.

Keywords: administrative data, quality, guidelines, integration

Title of abstract: Quality Measures and Indicators for Multisource Statistics

Ton de Waal, Arnout van Delden, Sander Scholtus
Statistics Netherlands

The ESSnet on Quality of Multisource Statistics is part of the ESS.VIP Admin Project. The main objectives of that project are (I) to improve the use of administrative data sources and (II) to support the quality assurance of the output produced using administrative sources. The ultimate aim of the ESSnet is to produce quality guidelines for National Statistics Institutes (NSIs) that are specific enough to be used in statistical production at those NSIs. The guidelines are expected to take the entire production chain in to account (input, process, and output). They also aim to cover the diversity of situations in which NSIs work as well as restrictions on data availability. The guidelines will list a variety of potential indicators/measures, indicate for each of them their applicability and in what situation it is preferred or not, and provide an ample set of examples of specific cases and decision making processes. Work package 3 of the ESSnet focuses on developing and testing quantitative measures and indicators for measuring the quality of output based on multiple data sources and on methods to compute such measures and indicators. Examples of such quality measures and indicators are bias and variance of the estimated output. Methods for computing these and other quality measures and indicators often depend on the specific data sources. Therefore we have identified several basic data configurations for the use of administrative data sources in combination with other sources, for which we proposes, revise and test quantitative measures and indicators for the accuracy and coherence of the output. In the presentation we will discuss the identified basic data configurations, quality measures and indicators for these basic data configurations, and methods to compute those measures and indicators. We will also point out topics for future research.

Keywords: administrative data, multi-source statistics, quality indicators, quality measures, survey data

Title of abstract: Quality Guidelines for Frames of Social Statistics

Thomas Burg, Magdalena Six
Statistcs Austria

Under the umbrella of the ESS.VIP Admin the ESSnet Kumuso dealing with quality aspects on multisource statistics was established. Within SGA 2 of this collaborative projecta work package was dedicated to draftingquality guidelines for frames in social statistics. The paper describes the motivation and the generation process of the guidelines and presents the document in its current state. After describing thestructure of the documentas an outcome of a stock taking actionbased onresults of a survey conducted among EU member states and investigations atNSIs outside the ESS regarding the construction, use and assessment of sampling frames within the NSIthe paper outlines the definition of frames in social statistics and the associated processes. In the second part the paper looks in more detail to the contents and substance of the guidelineextending the focus to other possible forms of using a frame then sampling such as input for processing and/or direct production of statistics. Another aspect which is considered in more detail is what kind of quality indicators and procedures are proposed to assess a frame in social statistics. Finally the possible present and future role and further development of the presented quality guidelines within the ESS and for NSIs is discussed.

Keywords: Quality Guidelines, Frames in Social Statistics, Multi Source Statistics

SPECIAL SESSION 38 – THE STATISTICS CODE OF PRACTICE FOR THE EUROPEAN NEIGHBOURHOOD POLICY (ENP) SOUTH COUNTRIES AND ITS IMPLEMENTATION – EXAMPLES AND PRACTICE

Title of abstract: Design of a roadmap for implementing the CoP for the ENP south countries at ONS-Algeria

Tarik Bourezgue
Office National des Statistiques, Algeria

The rapid changes in Algeria have had and continue to have a profound impact on infrastructure, economic agents and the population as a whole. These transformations have already turned the statistical landscape upside down and will continue to do so. In fact, the need for statistical data has changed and evolved in terms of the nature of the statistics, on the one hand, and requirements regarding availability, quality and time, on the other hand. Starting from the context described above, the development ofONS-Algeria's quality approach is essentially based on thecapitalization of the work of cooperation with Eurostat ontheprinciples of the European Quality Assurance Framework for Official Statistics(QAF), the option for aparticipatory and transparent processto enrich thisapproachand facilitate its appropriation and lastly, conducting the process in stages. In driving the process step by step, to optimize its management, our paper will address the issues related to the design of a roadmap for implementing the Code of Practice (CoP) for the ENP south countries.

Keywords: Code of Practice, implementation challenges, quality management, ENP South countries

Title of abstract: Implementation of the Tunisian Statistical Charter and linkage with the Statistics Code of Practice for the European Neighbourhood South countries

Mohsen Hajjej, Atef Ouni
National Institute of Statistics of Tunisia

After the revolution of 2011, Tunisia underwent major political, economic and social changes, including progress in terms of democratic transition with the adoption of a new Constitution that ensure the right of access to information. The National Statistical System is facing many new challenges, such as the increase in data demands at the national and regional levels, the improvement of data quality and the dissemination activities in accordance with international standards. A twining program started at the beginning of 2016 for two years to develop the NSS through a partnership with the European Union. Among the main objectives of this program are:

- The adoption of a new statistical law.
- The implementation of the Tunisian Statistical Charter (TSC), based on the other frameworks (African Charter on Statistics, Fundamental Principles of Official Statistics, CoP-ENP SC) and approved as a national framework organizing statistical activities, in order to provide information with quality that meet the user's needs.
- The revision of the organization chart of the INS-Tunisia, setting up of a methodology and Quality unit.

- A Roadmap for the preparation of the ISO certification.

The paper will review the context of the TSC and its linkage with the other frameworks (ACS, CoP-ENP SC, UN FPOS), as well as the contents of the TSC, focus on the components of implementation action plan. Will present also the challenges of TSC implementation, the work done for the establishment of the quality unit, the implementation of the ISO and GSBPM on the pilot operation will be also detailed.

Keywords: Code of practice, EPN south countries, quality management

Title of abstract: Implementation of the Statistics Code of Practice in Jordan (Good practices and vision)

Duraid Shawawreh, Abdel-Naser Tahat
The Department of Statistics in Jordan

The General Statistics Law No. (12) Of 2012 establishes the role and responsibilities of the Department of Statistics in Jordan.

The Law has been broadly effective in supporting the Department's independent statistical activity noting that the Department implements a range of activities in appropriate ways; the external stakeholders hold the Department of Statistics, including its de facto professional independence, in high regard. According to the law, ministries and institutions planning to conduct surveys should seek the advice and permission of the Department before they may do so. The coordination and cooperation role of it is clearly defined in the General Statistics Law of 2012. For most Statistical activities, the Department of Statistics coordinates through ad hoc committees that bring together users and producers in addition to other stakeholders. Moreover, any unofficial body that plans to collect and disseminate statistics on behalf of another body must obtain written permission from the Director General of the Department of Statistics. The statutory provisions for the protection of confidentiality by the Department of Statistics staff are consistent with the Statistics Code of Practice for the ENP-South countries. The Department of Statistics has achieved a high level in terms of technological progress and speed. It uses tablets (equipped with mobile chips) as a collection tool instead of the paper questionnaire. Our paper looks at these good practices as well as the vision of the Department of Statistics in relation to future work which is to draw a road map that ensures producing high quality statistical products.

Keywords: Code of Practice, The General Statistics Law, technological progress

Title of abstract: Importance of Implementing EFQM Model for Official Statics Institutes: PCBS Experience

Daliah Shahin
Palestinian Central Bureau of Statistics

This paper demonstrates how PCBS worked hard to fulfill the requirements of the certificate of "Committed to Excellent-C2E" delivered by EFQM, as a part of its efforts to develop comprehensive framework for the management of its operations according to international standards, this has been done to compete globally, and improve performance. The contribution of the paper is to explore an overview of EFQM model and the concept (EFQM model standards, Fundamental Concepts and RADAR Logic) and to highlight the concept of Excellence, Importance of applying EFQM model and the approach followed. It will also show the advan-

tages and benefits of applying EFQM model such as: Comprehensive framework for improvement, improvement of quality, optimizing resources and will explain difficulties and challenges faced PCBS in applying EFQM model. In addition to that, there is an overview of self-assessment process has been implemented in accordance with the European EFQM Excellence Model it had been successfully conducted based on the principles and standards of excellence derived from the European Model and its evaluation-by RADAR. Throughout this process, the current situation of PCBS was diagnosed in comparison with standards of excellence to identify gaps and determine aspects of opportunities as way to develop PCBS and its methodologies in order to achieve PCBS mission and goals; we will show the projects resulted from the self- assessment as opportunities for improvement. Finally the paper shows PCBS intention and its future steps towards continuing the way to achieve higher levels of excellence and sustainability.

Keywords: EFQM, Radar Logic, Committed to Excellent-C2E, Self-assessment

Title of abstract: Strengthening of data quality with limited resources: the case of the statistical database of the High Commission for Planning, Morocco

Mohamed Salimi
High Commission for Planning, Morocco

This paper will present an innovative project in the field of modernization of the process of the statistical production and dissemination with the objective of strengthening the data quality. It consists on the reverse of the approach used to collect, process, compile and disseminate statistical data. Concretely, instead of the current approach which adopts the following chronological order: data collection, treatment, dissemination and the update of the Statistical Data Base, the new approach is based on a new logic: Data collection, treatment, validation and update of SDB then automatic generation of dissemination products. To establish this model plan required a redesigning of our SDB by, among other improvements, building it on indicators-basis instead of series-basis and developing a module for generating automatically the publication "Morocco in figures" and an SDMX module for exchanging data on indicators, including MDGs, with the UNSD. The new tool will serve as well to exchange data automatically on SDG indicators. In addition, the new database offers the possibility to disseminate data on the intranet and the web for users. The paper will explain the problems faced, how they were fixed and the benefits in terms of data quality and in rationalizing used resources and discuss challenging specific objective to reach is to use SDMX for collecting data from the maximum of line ministries directly to the new SDB, completing the automation of the overall process in a context of reduced staff and resources and where ITs provide revolutionary solutions aiming to ameliorate the quality of the data.

Keywords: Automation, Data processing, Data Quality, Dissemination, SDMX

SPECIAL SESSION 39 – THE REVISED EUROPEAN STATISTICS CODE OF PRACTICE – CHALLENGES AND OPPORTUNITIES

Title of abstract: Special session: The revised European statistics Code of Practice – challenges and opportunities

Eduardo Barredo Capelot, Claudia Junker, Zsuzsanna Kovacs
Eurostat

The recent revision of the European statistics Code of Practice presented a challenge insofar that a good balance of two objectives had to be found: the objective to keep it as stable as possible over time, so that comparison is possible for measuring its implementation, on the one hand and to adjust it to new realities and challenges in developing, producing and disseminating European statistics, on the other hand. The paper will describe main elements of the revision process, such as the analysis of shortcomings of the Code of Practice adopted in 2011, basic considerations behind the changes, the consultation rounds and adoption. The main reasons for changing the Code will be described, including the need to emphasise a principle of and ways of coordinating the national statistical system, move away from a survey-based statistics production to a production based on multiple data sources and to reflect innovation in the Code. After the adoption of the revised Code of Practice, the way of measuring its implementation should be considered as well as the changes and opportunities that the revised Code of Practice may bring for other elements of the quality framework.

Keywords: Code of Practice, revision, implementation challenges

Title of abstract: Innovation and modernisation in the revision of the European statistics Code of Practice

Giorgio Alleva
Italian National Institute of Statistics (ISTAT)

The external world is rapidly changing, producing dramatic shifts, both in the statistical process, the production model and in its output. The producers of official statistics face new challenges. The growing availability of data discloses new opportunities, in particular about the use of administrative archives, the adoption of increasingly accessible unstructured sources and integration of data from multiple data sources. But at the same time it becomes crucial for the NSIs to have access to administrative data and to rethink and to extent the role of sampling surveys in a different data environment. The coordination role of the NSIs represents a strategic key aspect in a modernization process which is advancing in the ESS as a whole, as well as the cooperation within the European and National statistical system and beyond finding new ways of working together in order to fulfil both European and national statistical needs. New data sources, as Big data, require that we rethink our traditional statistical tools, review our IT architectural infrastructure, update our quality commitment looking at innovation, and enhance research and partnership among other NSIs and other producers. The ESS has long-term service capability to adapt fast enough to keep up with changes in the environment, but if National statistical institutes want to produce trustworthy data from new data sources, they must introduce concepts and tools for quality evaluation specifically devoted to a multisource environment. The European Code of practice has been revised including key aspects mirroring this new challenges.

Title of abstract: Working towards building a robust implementation technology for the revised Code of Practice

Athanasios C. Thanopoulos
Hellenic Statistical Authority (ELSTAT)

The recent revision of the CoP was the result of an effort aiming to adjust and/or fine-tune the existing text of the Code to new facts related to the data revolution and to the forthcoming challenges stemming thereof. This has to be coupled with a solid implementation and monitoring framework through tools that would measure implementation of the CoP in an objective, uniform, coordinated and transparent manner, in order to fully benefit from the momentum of the recent revision. The main channel used so far for the implementation of the CoP was the incorporation to the national legal order. Typically, this would take either (a) a direct form, i.e., mot-a-mot transposition of the text of the CoP into the national Statistical Law, or, (b) an indirect form, i.e., through legal instruments such as Commitments on Confidence in Statistics (CoC). There are advantages and disadvantages associated with either form. Most salient drawback of (a) is the fact the CoP is not meant to be directly used as stricto sensu legal text, while the most apparent shortcoming of (b) is the fact that it requires updating and adaptation, it may have become (partly) obsolete or irrelevant. Given the fact that neither legal texts regulating Official Statistics are meant to be altered with a high frequency, nor is it the case that the CoP gets revised in a pre-scheduled and linear manner (with respect to time and with respect to the importance of its revised text), this paper should address the following questions:

(a) Is the above implementation (and monitoring) technology still relevant in light of the revised code?

(b) Have there been identified inherent inconsistencies/ conflicts that weaken the effectiveness of this implementation framework?

(c) What are the desired properties that an implementation / monitoring technology should have in order to be robust to future revisions of the CoP?

Keywords: Code of Practice

Title of abstract: Quality guidelines as a tool for ensuring the coordination of quality in the French statistical system

Christine Fluxa
INSEE, Paris, France

The revised code of practice includes a new principle relating to coordination and cooperation (principle 1bis); more particularly, according to the indicator 1bis.2, NSI have to produce national quality guidelines to ensure the coordination of quality in their national statistical system. In compliance with the amended European regulation 223/2009 the French Quality Unit established in early 2017 quality guidelines in collaboration with the ministerial statistical services. These guidelines ensure that the head of the French NSI has access to indicators enabling to monitor the quality of the most important statistics produced in the public statistical system. These guidelines formalize the framework for the co-ordination of the statistical system in terms of quality. They must be applied to European statistics, in accordance with the code of practice. But they also concern the national statistics that have been described as «structuring». These «structuring» statistics are defined as both highly expected by the users and extremely prejudicial to the services if they were

to be of poor quality. Their production and dissemination need to be insured against risks. These guidelines are based on five orientations concerning governance, development of quality skills within the statistical services, the planning and realization of quality approaches to statistical processes; the fulfillment of European commitments and the systematic integration of users' needs and satisfaction in the designing of statistical products and services. This article firstly describes the French quality guidelines and how they are initiating the management and the integration of quality in statistical processes. Then, it presents the procedure that has been followed in order to support each ministerial statistical services in the implementation of the guidelines. The article lastly shows how this implementation is monitored by the French NSI and how it will help to fulfill our commitments resulting from the 2014 peer review and to co-ordinate quality issues among the French statistical system.

Keywords: Co-ordination, National statistical system, guidelines, institutional environment, ONA, governance, revision of the European statistics code of practice

Title of abstract: The Revised European Code of Practice – a new opportunity for European statistics

Enrico Giovannini
European Statistical Governance Advisory Board

The European Statistics Code of Practice was established in 2005 and sets the standards for developing, producing and disseminating European statistics. It aims to improve confidence in statistical authorities by proposing specific institutional and organisational arrangements which could be put in place. It also aims to reinforce the quality of European statistics by promoting the application of best international statistical principles, methods and practices. The Code is updated from time to time in order to keep up with changes in statistical legislation and statistical practices. The first update was in September 2011 and the second was in September 2017. The 2017 update takes into account new legal requirements concerning the coordination of statistical activities in national statistical systems, and issues such as access to new types of data, the use of multiple data sources, data protection, and the use of open data standards. This paper discusses the reasoning behind the key revisions to the Code and the opportunities they offer for improving statistical governance. The paper also considers the implications of the revisions with regard to the process of monitoring the implementation of the Code.

Keywords: Code of Practice, governance

SPECIAL SESSION 40 – BETTER LOCALISATION, BETTER QUALITY – THE GEOSPATIAL DIMENSION OF QUALITY IN STATISTICS

Title of abstract: Integration of GEOSPATIAL DATA within the statistical production process – GEOINQ

Ana Santos, Mário Lucas, Francisco Caldeira, Bartholomeus Schoenmakers
Statistics Portugal

The use of cartography has supported census data collection at Statistics Portugal since 1981.

Since 2006, with the production of the enumeration areas dataset to support the 2011 census, Statistics Portugal has been developing a Spatial Data Infrastructure (SDI) that is currently being used, in a transversal way, to promote the integration of the geospatial data in the statistical production process, in order to achieve efficiency and accuracy, within sampling process, data collection or dissemination. In 2011 Statistics Portugal built a national geodatabase comprising all the georeferenced buildings from the 2011 Census (BGE). BGE is a point based coverage that is being continuously edited in a internal quality control process and updated by the Municipalities which provide Statistics Portugal, on a monthly basis, all the completed buildings and buildings permits including X,Y location and addresses. After the 2011 Census, Statistics Portugal evaluated the possibility of implementing a geographic tool that would allow the visualization of buildings and dwellings units of survey samples. For this purpose, GeoINQ was developed to allow the visualization of the location of sample buildings and to provide management and control functionalities of the data collection process. In addition, GeoINQ has editing tools to update the BGE by field interviewers.

It is a Geographic Information Systems (GIS) WEB solution developed in order to integrate geospatial data into the production process of official statistics in an innovative way. It allows greater efficiency and rationalization of the resources especially in the household's surveys by supporting the data collection process. GeoINQ is integrated, via webservices, with Statistics Portugal Global Survey Management System (SIGINQ--IE), and consumes a set of services and geographic data of subjects of the INSPIRE directive.

Keywords: GEOINQ, Geospatial data, Webservices, Surveys, Statistics

Title of abstract: Improving the quality with spatial sampling

Vincent Loonis
INSEE, Paris, France

Over the last years, many initiatives have been undertaken to help the National Statistical Institutes construct a fully geocoded information system. Such a point-based system is precisely the starting point of a handbook of spatial statistics being written by Insee and to be released by the end of June 2018. The latter, funded by Eurostat, will drawa list of statistical methods that rely on theavailabiltyof the (x,y) coordinates of the statistical units. Ranging frommeasures of spatial autocorrelation to spatial econometricsfor panel data, these methods might be helpful to improve the production, the dissemination or the analysis of statistical results. The issue of spatial sampling fully falls within the scope of the handbook, and as such will be more

precisely dealt. The presentation at the conference will focus on the issue of spatial sampling. It aims at proving that a geocoded sampling frame might help better carry out surveys. On the one hand, knowing the position of the statistical units help better organise the field work for face-to-face surveys. This can be done, for instance, thanks to Primary Units having very good spatial features. On the other hand, at the selection level, the sampling design might better spread the selected units over the territory. This strategy can be efficient to improve the precision estimation for high spatial correlated variables.

Keywords: Integration of geospatial and statistical information, geocoded sampling frame, spatial sampling

Title of abstract: Making census statistics more relevant – towards geo-enabled statistics

Adam Wroński, David Thorogood
Eurostat

At a time when there is evidence of public distrust of statistics and of distancing between policy makers and the public, it is important that official data should relate closely to the realities of people's lives. One way to do this is to provide geographically detailed statistics that describe the situation at local level. In many contexts, statistics at the level of towns and neighbourhoods are the most relevant and most meaningful to the citizens and local authorities to support the local level policy decisions that are of importance to people's day-to-day lives. For the 2021EUcensus, building on the success of the GEOSTAT2011 grid and in response to growing needs for high resolution population statistics, the work is advanced on the 13 key census topics geocoded to a 1km^2 grid. It will be a new departure, allowing new types of statistical and spatial analyses for areas that unlike the traditional NUTS areas can be flexibly defined according to the needs of policy makers and researchers. Grids as a time series will also assure geographical comparability over time. One use of such data to ensure the efficient allocation of funds (e.g., European Cohesion Policy €50 billion/year) is the analysis of populations for small urban and rural areas, and assessing the accessibility of services such as education and health for citizens. This paper presents the current plans for developing the grid data collection for the 2021-census and discusses the main legal, organisational and methodological issuesto be addressed. It describes the innovative and complex development, requiring a separate legal act and a number of solutions to emerging technical and methodological issues, e.g. Statistical Disclosure Control. Itlooks forward towards the post-2021-EUcensuses as the attractiveness of providing population data to a grid made it one of the focal points of the developments.

Keywords: population census, geo-coding, spatial analyses

Title of abstract: Towards better quality of statistical geospatial data production – harmonization of statistical and geodetic divisions in the context of the 10 Level Model

Anna Sławińska, Janusz Dygaszewicz
Statistics Poland

According to adopted rules, in Poland boundaries of statistical units are aligned with boundaries of cadastral units where applicable. In connection with the need to preserve the limits of housing and people in the statistical units and taking into account the diversity of terrain and population density, statistical regions and census enumeration areas in rural and urban areas have different extend. To preserve collinearity in case of any changes made by the cadastral service appropriate changes must be made on the statistical side. In reference to keeping a consistency of both divisions Polish proposal of the 10 Level Model will be used to better understand and develop statistical and geodetic reference framework. The proposed model should be the subject of intensive works in order to overcome existing barriers and as a starting point to make practical progress in the methodology of combining spatial data with statistical data. Development of statistical division based on geodetic division should provide better interoperability of sets of data and raise possibilities of statistical geospatial analyses. However such harmonization will cause that quality of statistical geospatial data as well final statistical products and conducted analyses will depend on the quality of input spatial data from external administrative registers. That is why assessment of the overall quality of external spatial data sets, and especially the quality of data which they include is essential. Recently Polish official statistics worked on the project which aim was the improvement of the use of administrative sources. As a result, on the basis of Polish experience, the methodology of assessing the usability of administrative data sources including spatial data registers has been elaborated.

Keywords: geospatial registers, quality of geospatial data, quality assessment, statistical and geospatial data integration

Title of abstract: Accessibility statistics and data integration: from remoteness and public transport to cultural accessibility

Pasi Piela
Statistics Finland

Accessibility research has been a relevant part of today's geographic information science. Data sources of official statistics offer plenty of relevant administrative data of the population itself and of many kinds of services it is potentially using. These are also combined to statistical products for customers of Statistics Finland. This paper discusses many aspects of accessibility statistics. The target here is the whole country. First, the remoteness (index) estimation that has been conventionally computed by Euclidean buffer populations for the Ministry of Finance. Studies funded by the ministry clearly promote the use of a road network based estimation. Elementary school accessibility has been taken into account as well. These kinds of statistics at the municipal level offer valuable information for the relevant authorities. Recently, the need for cultural accessibility statistics has been raised at Statistics Finland. This requires data from many sources: libraries, theatres, movie theatres, orchestras, museums, festivals, etc. Naturally, the road network application is much more suitable than linear distances. The most advanced application is commuting statistics. Here, the travel time itself becomes relevant and estimating that during the rush hour is a complicated task requiring data integration of many different sources. On the other hand, new data sources such as public transport web services enrich the modelling much. Greener commuting, bicycling, has received the most attention, however. This paper both informs and motivates for discussions of current and future research of statistical accessibility.

Keywords: Accessibility, network analysis, data integration, GIS

SPEED TALK SESSION 01 – QUALITY MANAGEMENT AND FRAMEWORKS, QUALITY ASSURANCE, COSTS

Title of abstract: A common quality framework for Swiss federal statistical offices

Markus Baumann
Swiss Federal Statistical Office

The presentation proposed by Switzerland concerns the creation of common quality criteria that can be used by all swiss federal statistical authorities. It aims to present the various stages and difficulties encountered so far, the defined quality criteria and the possible mechanisms for their implementation and evaluation in the future. Having a coherent and effective national statistical system has been a priority for the members of the European Statistical System for many years. Switzerland is no exception to this reality and the Federal Statistical Office is committed to intensifying the dialogue with its federal partners on quality in public statistics. The notion of „quality" is undoubtedly complex and multidimensional. A working group has been set up bringing together several actors from the swiss federal statistical system to define a set of concrete and comprehensible quality criteria. In order to do so, the working group relied on the European Statistics Code of Practice (CoP), the recommendations of the Peer Review 2014/2015 of the swiss statistical system and the "ESS Quality Guidelines and Performance Indicators" but has limited its reflections on the elements that have emerged as priorities. This made it possible to have a narrower reference framework than that of the CoP, while ensuring an acceptable level of compliance for users and encouraging their implementation and continuous improvement.

Title of abstract: The implementation of the OECD Recommendation on Good Statistical Practice: Professional independence and coordination

Julien Dupont
OECD, Paris, France

In 2015, the Organisation for Economic Development and Cooperation (OECD) adopted its first legal instrument for statistics, i.e. the Recommendation on Good Statistical Practice (hereafter the Recommendation) intended to provide a common reference to assess the quality of national statistical system and official statistics, which are fundamental for OECD statistical and analytical works. The Recommendation complements existing codes of practice and international standards currently applied by OECD Member countries, but is also more specifically relevant for OECD statistical activities, for example in including good practices on the use of new sources of statistical information, or on the coordination of the statistical system.

The Recommendation includes twelve recommendations structured in five main areas: (I)institutional, legal and resource requirements that enable statistical systems to function; (II)methods, quality and processes of statistical production; (III) dissemination; (IV) co-ordination and co-operation; and (V) statistical innovation. Each of the twelve recommendations is presented with a set of indicative good statistical practices based on the OECD's experience in statistical reviews.

This paper presents the Recommendation and provides an overview on the importance of quality for an international organisation as the OECD. While the professional independence of national statistical au-

thorities was questioned in several recent cases in OECD member and non-member countries, the paper also sheds some light on specific questions related to this principle as well as to the coordination of the National Statistical System. To this end, preliminary results from the activities currently carried out by the OECD in order to assist adherents in implementing the Recommendation.

Keywords: National Statistical System, professional independence, coordination, quality

Title of abstract: How to turn quality into a habit in the statistical production?

António Agostinho, Paula Silva, Margarida Pinto
Banco de Portugal

One of the main purposes of the Statistics Department of Banco de Portugal is to ensure a statistical production with high quality standards aiming at fully meeting users' needs, aligned with the best practices and procedures recommended by the international organizations. Following its commitment to quality, one of the Bank's priorities is to develop a wide set of quality control procedures that ensure high levels of regular and thorough review of the key statistical outputs. Statistical quality control is based on different procedures and working arrangements that make sure that processes are effective and efficient and the risks are mitigated. In order to achieve higher quality statistics, there are several quality indicators performed by the primary statistics' compilers. This paper will present the main quality indicators used and the ongoing process to improve the model of regular and systematic quality controls.

Keywords: statistical quality control, quality assessment, quality indicators

Title of abstract: A Framework for Assessing the Quality of Banking Supervision Data

Gregorio Guidi, Francesco Donat, Iulia Tintea, Luis Antonio Suarez Tumi
European Central Bank, Frankfurt am Main, Germany

Since the start of the economic and monetary Union, the European Central Bank (ECB) has placed strong emphasis on key aspects of statistical quality, as described in its Statistic Quality Framework. The introduction of a harmonised framework for supervisory reporting in 2014 has raised the awareness of competent authorities on the need to apply similar high standards of statistical quality to data collected for the purpose of banking supervision. Accordingly, the ECB has established a process for the quality assessment of supervisory data, with a twofold purpose. Firstly, checking whether data constitute a suitable basis for informing supervisory decision; secondly, recognising data quality as an integral part of banks' supervisory evaluations. Following these two objectives, a number of data quality dimensions have been identified and implemented as part of a general framework for assessing the quality of banking supervision data. This paper presents some reflections on the framework, based on the experience gained at the ECB over the past four years.

Keywords: Banking supervision; BCBS 239; Data quality; ECB; Supervisory reporting; Statistics Quality Framework

Title of abstract: Quality Assurance and the GSBPM adoption at INEGI

Gloria Martha Rubio Soto, Enrique Ordaz, Eduardo Jallath
INEGI, Mexico

At the end of 2014, INEGI introduced new institutional and technical measures for strengthening data quality, as part of the Quality Assurance Norm (QAN). The Norm outlines a general framework for quality assurance, establishes assessment requirements and defines institutional arrangements. The QAN implementation will involve several phases over the medium and long term. The initial implementation phase comprised the quality framework definition, a pilot self-assessment exercise, and priority setting. More recently, the quality assurance reform has been intertwined with the gradual adoption of the General Statistical Business Process Model (GSBPM). The document will review the progress to date, discuss the lessons learnt from the early implementation experience and examine the challenges ahead.

Keywords: quality framework, quality assessment, General Statistical Business Process Model

Title of abstract: Cost Accounting of Products in Turkish Official Statistical Program

Deniz Ozkan, Nesibe Gul
Turkish Statistical Institute

Cost accounting is a relatively new concept in the public administration. Many countries have adopted this approach since the 1990s. It helps increase productivity by providing essential inputs to decision makers. But it is not a common practice among National Statistics Offices to measure the cost by statistical product. International organizations are supporting cost analysis of statistical products. United Nations' Sustainable Development Goal indicator 17.19.1 is about the measurement of "dollar value of all resources made available to strengthen statistical capacity in developing countries". In 2016 and 2017 TurkStat conducted two studies to calculate the cost of statistical products in Turkish Official Statistical Program. One was done to calculate the cost of statistics produced by TurkStat and the other to calculate the cost of statistics produced by other national authorities (ONAs). First, we measured costs of all statistics produced by TurkStat for 2015 and 2016. In this study both direct and indirect costs were calculated by product. Indirect costs are allocated to statistics by cost drivers. Cost driver must be a kind of factor which has the highest impact on all cost components. The number of personnel and their salaries are used as cost driver; since personnel expenditure makes up the largest part of total cost. Second, a survey was undertaken with ONAs to collect data on the number of personnel working on the production of statistics and their full time equivalent in 2015. Personnel expenditure was calculated by statistics. However indirect costs were not included in the cost calculation for ONAs. Methodology and results of the cost analysis study in Turkish Official Statistical Program are presented in this paper.

Keywords: cost analysis, statistics, full time equivalent, cost driver, direct cost, indirect cost

Title of abstract: The NSO, the NSS and Beyond!

Laurie Reedman – Statistics Canada
Marsha Windross – Statistical Institute of Jamaica

This paper takes a brief look at some early ideas and concepts about quality management at a National Statistical Office (NSO), and reflects on the breadth and depth of content included through extensive collaboration in the first generic National Quality Assurance Framework (NQAF). Since the deployment of the generic NQAF in 2013, many NSOs have adapted it to their own circumstances and adopted either a regional or their own national version. The journey is long yet fruitful. The exercise of reviewing priorities, challenges, bottlenecks and inefficient practices provides an opportunity to develop and document good quality practices, and goals to work towards. While the expected audience of an NQAF is the NSO, in few countries does the NSO produce all or even most of the official statistics. Hence the intended audience for an NQAF should really be all federal ministries, departments and agencies (MDAs) participating in the National Statistical System (NSS). The latter half of this paper looks at the experiences of the Statistical Institute of Jamaica providing data quality workshops to introduce other participants of the Jamaican NSS to quality assurance, and the experiences of Statistics Canada in producing a data quality toolkit intended for data producers and users outside of the NSO.

Keywords: Quality Assurance Framework, data producer, quality practices

Title of abstract: Improvements on the Brazilian Statistics Code of Practice

Maria Luiza Barcellos Zacharias, Raquel Rose Silva Correia
Brazilian Institute of Geography and Statistics, IBGE

In 2013, the Brazilian Institute of Geography and Statistics (IBGE) made available its Statistics Code of Practice, disseminating a set of guidelines, principles and practices that the Institute is committed to uphold in the statistical production process, taking as reference the Statistics Code of Practice for Latin America and the Caribbean. The Code aims at standardizing professional procedures to foster best practices in statistics, which are crucial to establish institutional credibility and, consequently, trust in the information produced by the Institution. The Code sets out 17 key principles and 80 good practices, concerning the institutional environment and coordination, and the statistical processes and products. In 2016, the Institute underwent an external auditing to assess its compliance with the principles and practices of the Code of Practice, which resulted in a set of recommendations that led to its revision and improvement. One recommendation was to develop and incorporate explanatory notes that clearly outline the context and objective of each good practice, in order to reduce the possibilities of different interpretations. Furthermore, a set of measurable criteria (quantitative or qualitative indicators) should be developed and incorporated to the Code as a means of assessing compliance with each practice. This paper presents the steps developed, the external references and frameworks used, and the choices that were made during the revision process. The IBGE Statistics Code of Practice is also seen as a mechanism to introduce a common understanding of quality across the producers of official statistics in Brazil and align national practices with international standards. The new edition of the Code will provide the basis for a more comprehensive version to be discussed with other institutions responsible for the production of official statistics at the National Statistical System.

Keywords: Code of Practice, Quality

SPEED TALK SESSION 02 – QUALITY MEASUREMENT, ASSESSMENT, REPORTING, METADATA

Title of abstract: But are those numbers correct?: Towards criteria to assess the reliability of statistics

Simon Scott
OECD, Paris, France

Confronted with a new statistic for the first time, how do we know whether to trust it?

Since 1990, several international documents have enunciated principles and practices for ensuring sound statistics. But they do not spell out the inherent attributes of sound statistics. Such attributes can be found in "quality frameworks", but these often include practical aspects such as accessibility and timeliness that say nothing about whether data are actually correct.

A checklist for assessing the reliability of data could have multiple uses. In particular, it would help establish whether a given statistic should be accepted as knowledge, and thus as suitable for use in hypotheses, or as a sound basis for action or policy. It might also help test whether existing agreed pinciples suffice to ensure reliable data. This paper is a first attempt to identify characteristics of data that may be accepted as true.

The analysis is in three sections, corresponding to stages of statistical production, namely:

Measurability: this section introduces the idea of an "evidence threshold" separating knowledge from speculation. It emphasises, however, that statistical quantities are found along a continuum ranging from known countable items to model-based projections.

Measure: this identifies characteristics of measures that promote reliability, including conceptual clarity, definitional rigour, and unambiguous application to the target variable.

Measurement: this stresses the value of comprehensive counts, adequate and competent enumerators, the absence of incentives to distort figures, and mimimisation of processing steps.

The discussion draws on official and non-official critiques of statistical quality, highlights common errors, and cites practical examples. It aims to initiate discussion of general criteria to help users assess the credibility of statistics, and help compilers maximise the reliability of their output.

Keywords: Knowledge, reliability, truth, evidence, measurability

Title of abstract: Comparison of datasets as a method of monitoring data quality: A conceptual framework and two applications with banking data

Angelos Vouldis, Stefano Borgioli, Martina Spaggiari
European Central Bank, Frankfurt am Main, Germany

Cross-validation using different datasets reporting comparable data can be a very effective way of monitoring data quality of macro data and complements validation rules applied within a single dataset. In particular, cross-validation can identify data quality issues which would not be found with the validation rules approach, specifically those related to the completeness or double-counting of information. This approach is especially relevant for banking data, given that, due to the need of informing different types of policies related to the banking sector, like monetary, macro-prudential and micro-prudential policies, various banking datasets have become available at central banks and banking supervisors. These datasets differ with respect to their scope, method of consolidation and granularity. They also differ in their collection method and the institution which processes the information. Despite these differences, one can formulate logical relationship which should hold between the different aggregates (e.g. at a jurisdiction level or at a bank sector level, like 'Significant Institutions'). Two examples are then provided. First, a comparison between the Consolidated Banking Statistics (CBD) which are used mainly for macro-prudential analysis and the Supervisory Banking Statistics which are collected in the context of the banking supervision exercised by the Single Supervisory Mechanism at firm-level. Second, we investigate the international interconnectedness of banking systems in the International Banking Statistics of the Bank for International Settlements and the CBD dataset of the European Central Bank. In both cases, we find that these comparisons provide important insights about the underlying data and allow the identification of data quality issues which could not be achieved through intra-dataset validation rules.

Keywords: Quality Assessment, Validation levels 4 and 5, Banking data, Macro-prudential policy

Title of abstract: Supporting the compilation of quality reports – Improvement of guidelines, provision of a checklist and standard texts

Irina Meinke
Federal Statistical Office (Destatis)

During the last years, the quality unit of Destatis made valuable experiences at national and international level with the compilation of quality reports concerning the following questions:

- Which concepts of the quality reports are typically posing problems for the subject matter units?
- How can existing guidelines be improved (in wording and form) in order to better support the compilation of quality reports? Based on the experiences made by Destatis, the aim of the paper is to present which additional support could be provided to the compilers of quality reports – besides the already existing ESS or national guidelines for quality reports:
- A checklist for quality reports based on the guidelines for quality reporting,
- Extensions and further specifications on the content of the guidelines for quality reporting,
- Provision of standard texts for designated concepts.

Keywords: quality report, ESQRS, ESMS, metadata

Title of abstract: Quality Reporting Improvement Depending on the Generic Statistical Business Process Model (GSBPM): PCBS Experience

Basila Samara, Hiba Masoud
Palestinian Central Bureau of Statistics (PCBS)

PCBS has been working on preparing reports about data quality of its statistical surveys, aiming at creating a general perspective regarding the extent of applying the quality indicators in statistical surveys. The contribution of this paper is to address three parts of quality report based on the Palestinian experience at PCBS:

The first part focuses on the historical overview of PCBS and Quality Department in the preparation of quality reports for statistical surveys. Starting with quality reports, they are prepared after the completion of the statistical survey and then we have developed and improved „operations and data quality reports" to control the quality during survey implementation, which contains many indicators associated with data quality in line with the GSBPM standard, and its cover many of process and sub-process in each phase (specifying needs, designs, building, collection, processes, analysis, disseminates, and evaluations in all phases), these reports help us to decrease the non-sampling errors.

Secondly, this paper also focuses on operations and data quality reports and their contribution in solving the problems that may face the project management, as well as reducing non-sampling errors; leading to the improvement of the data quality of statistical survey. In addition, the possibility to evaluate the quality of the survey during its different phases and it helps in determining the most important strengths and opportunities for improvement for each phase, through drafting recommendations to improve quality during the current or the next survey cycle, and documenting the results of monitoring data quality during the project cycle.

Finally, we are looking to improve the quality report by utilizing measurable quality dimensions depending on GSBPM.

Keywords: GSBPM, Non-Sampling Errors, Data Quality, Quality Indicator

Title of abstract: Structural metadata as a key element in the management of microdata

Ana Isabel Sánchez-Luengo Murcia
National Statistics Institute of Spain (INE)

Within the international framework of statistical production standards (GSBPM and GSIM) there is a clear aim of building a modern production system metadata driven. INE Spain is putting its efforts into an approach based in two main ideas; on the one hand, to create or reuse the metadata at the beginning of the process and maintain them throughout the process; and, on the other hand, to design structural metadata elements as GSIM objects. The model proposed in this paper uses structural metadata, i.e., variables, classifications, concepts and statistical units as an information system, key for the storage of microdata. These metadata are stored once in a single institutional repository and should be used and reused as much as possible. Designed and developed, the model is beginning to be implemented at INE-Spain. The structure metadata model allows to build microdata bases and access them easily. The access can be made to data with different temporal references and in different subprocesses of the data collection phase. As regards the

quality of metadata, a strategy has been put in place where the metadata unit is responsible for building, maintaining and improving structural metadata, applying international standards as much as possible, always bearing in mind the institutional needs.

Keywords: Structural metadata, Microdata, Statistical Process

Title of abstract: Measuring response burden at the Swiss Federal Statistical Office

Desislava Nedyalkova, Lionel Qualité
Swiss Federal Statistical Office

The Swiss Federal Statistical Office (SFSO) is willing and needs to put in place a procedure for systematically measuring the burden that stems from its different surveys (household and business surveys). A first evaluation of survey burden, in terms of time and cost imposed on businesses, was carried out in 2013 by external specialists at the University of St.Gallen, Switzerland. A working group within SFSO has recently examined the existing literature and the dispositions taken by other national statistical institutes. Its current task is to expand and improve on the 2013 works and thus define the way in which burden will be accounted for. This paper presents an overview of the difficulties encountered when measuring actual response burden, the principal challenges that need to be met and the different options available. It lists measures that have already been adopted so as to reduce the global survey burden, such as, for instance, mandatory use of administrative data, online questionnaires, profiling of large and multi-establishment companies. Also is to mention SFSO's tool for coordinated selection of samples, which allows to spread the response burden for most of our surveys. Among other benefits, it allows to track unit selections over time and thus facilitates survey burden evaluation.

Keywords: Response burden, Sample coordination

Title of abstract: Reduction of response burden by utilising extensively register data and modelling: Cases from new EU data needs in agricultural statistics

Esa Katajamäki, Johanna Laiho-Kauranne, Pasi Mattila, Anneli Partala
Natural Resources Institute Finland, Finland

In ESS statistics, the use of administrative data is supported, which is also emphasized in the ESS Strategy for Agricultural Statistics for 2020 and beyond. Natural Resources Institute Finland (Luke) has a long experience of using administrative registers as a source for statistics. However, when new data needs are expressed or the administrative data develops, new possibilities arise. Our objective is to replace survey data in forthcoming farm surveys with register data combined with advanced modelling in survey estimation procedures. We examine and demonstrate the recent advances accomplished in the usability of the agricultural registers data and other data sources to reduce both the response burden and direct data collection from the farms.

The case studies we examine are:

1) Greening the agricultural production and crop rotation.

2) Agricultural labour

3) Number of animals.

The general objectives in our study are:

1) To investigate possibilities for a broader analysis of crop rotation based on the IACS parcel data icluding the new geospatial parcel data obtained from the farmers through farm subsidy administration from the year 2015 on.

2) To pilot the use and examine the quality of the individual level register information on farm labour that can be linked with identificators to farm level from a register other than IACS as a source of FSS data.

3) To evaluate the quality and feasibility of all existing animal registers as sources of data for animal statistics.

4) To examine thepossibilities to respond to the new ad hoc information needs.

Common advantages of the use of registers are the total coverage of farms, reduction of survey costs and the avoidance of misinterpretations by farmers when answering the questionnaires which is a significant factor in the case of the crop rotation variable.

Keywords: renewing agricultural statistics, new information needs, administrative data, estimation,response burden, linkage information

SPEED TALK SESSION 03 – USE OF ADMINISTRATIVE DATA SOURCES, DATA INTEGRATION

Title of abstract: Integration of New Administrative Data Sources into Turkish Statistical System

Bilal Kurban, Mahmut Ozturk, Mustafa Arda Yilmaz, Hatice Burcu Eskici, Huseyin Tancan Kale, Serkan Arslanoglu
Turkish Statistical Institute

Turkstat aims to collect and process data more quickly, more accurately and more effectively than before. On the other hand; unlike the past, it seems like data collection works have been changing rapidly in recent years. Emerging technologies, respondent burden concerns, increasing costs, globalization, having more difficulties in collecting data than the old times and exponentially increasing amount of data day by day and thus increasing data demands are among the reasons for these changes. Overcoming the data collection challenges, Turkstat has given priority to „increasing the weight of administrative registers in statistical production" in its future plans and programs and has accelerated the studies on this subject. Turkstat started to pay more attention to take advantage of administrative sources. In this context, TurkStat started regular monthly data transfers from Revenue Administration as of June 2017 and from Social Security Institution as of October 2017. Starting from 2018, as well as structural business statistics characteristics, short term business statistics indicators such as turnover, production and labor input are planned to be produced by using these administrative records. Furthermore, as data collection shifts to secondary sources due to integration of new data sources, Turkstat started to have a new employee profile having combined capabilities of statisticians and IT specialists. Making use of new administrative data collection activities in Turkstat also started to cause developing new skills, creating a mindset and culture open to change and developing strategic partnerships with administrative registers holding public institutions. This paper shows Turkstat's experiences on integration of new administrative data sources into Turkish Statistical System, the resulting necessity of redesigning its business statistics and analytical skills and competencies developed in handling large and complex administrative datasets.

Keywords: administrative data, data integration, business statistics

Title of abstract: Managing changes in key administrative data sources within Finnish SBR

Kennet Härmälä, Antti Lempinen, Noora Seppä, Johanna Sisto, Ville-Matti Pilviö
Statistics Finland

The latest Eurostat Code of Practice (CoP) peer review was organized by Eurostat in 2013–2015. The Finnish country report (2015) identified, among other findings, a strong reliance on administrative data sources in statistical production. Recommendation #6 of the report called for risk assessment of possible changes in these administrative sources. The objective of the so-called Valmis-project (2013–2019) is to develop operations and reform taxation software. Numerous taxation software will be replaced by one software package (GenTax). These changes will affect the administrative data that the National Board of Taxes provides to Statistics Finland. There will be changes to both the content and the format of the data. At the end of 2015

the Administrative Data Collection team (AVa), part of the Data Collection Unit at Statistics Finland, prepared a report for The (National) Statistical Law Working Group outlining risks and risk management with the use of administrative data sources in statistical production. Key points of this paper include:

- An outline of the Tax administration's Valmis project.
- Analysis of the effects the Valmis project on The Statistical Business Registry (YTY).
- Assessment of the centralized input of data and change management procedures of YTY in adapting to these changes.
- Assessment of the benefits gained from risks management plans for identified administrative data sources.
- Analysis of methods to benefit quality control from materialized risk scenarios.
- Lessons learned from a materialized risk situation: a uniform abstract layer between data input and the Statistical Business Registry.

Keywords: administrative data, risk management, risk assessment, quality control, taxation software, GenTax

Title of abstract: Harmonisation across the UK: Comparability of survey and administrative data in conjunction with European influences

Charlotte Wroth-Smith, Daisie Hutchinson
Office for National Statistics, United Kingdom

The Harmonisation Team work across the UK Government Statistical Service (GSS) by harmonising definitions, questions and outputs to ensure that the quality of official statistics meets user requirements and supports comparability both within the UK and internationally. The GSS harmonisation vision is that all definitions, questions and outputs for the census and surveys and all data from administrative records will be harmonised, so that users can compare data from different sources with confidence and can merge and match data more easily. The harmonisation of administrative data is a new and challenging area of work and forms a large part of this harmonisation work programme. Understanding the conceptual differences between survey data and administrative data is key to this. The harmonisation team plan to develop a set of harmonised definitions, questions and outputs for administrative data for use across the GSS. This will be in addition to the harmonised definitions, questions and outputs already in place for surveys. A further new area of work is the harmonisation of business statistics to comply with EUROSTATs Framework Regulation Integrating Business Statistics (FRIBS). The regulation requires the GSS to move to a harmonised set of variables by 2019. Alongside this, and in collaboration, ONS are harmonising business survey questions where possible on the Electronic Data Collection programme over the next few years. The paper sets out what the GSS has achieved to date with harmonisation, including the development of a harmonised question library, harmonised definitions, questions and outputs for surveys, administrative data and also what remains to be done. It also outlines the benefits of harmonising and details the issues and challenges faced when attempting to harmonise.

Keywords: Harmonisation, comparability, quality

Title of abstract: Record linkage in agricultural statistics

Anders Grönvall
Swedish Board of Agriculture

The use of data from administrative registers have been used extensively in Sweden since Sweden became a part of the European Union in 1995. Integrating administrative registers with censuses and sample-surveys has been seen as a cost-effective way of producing statistics with sufficient quality. The integration phase where data from several sources is integrated into a new statistical register is seen as essential for achieving sufficient quality. To successfully link records from a specific unit in an administrative registers with a corresponding unit in a statistical register is therefore essential for the quality of the final statistics. In some cases, the linkage is perfect but in many cases, the unit in the administrative record do not uniquely relate to the unit in the statistical register. Choices and rules taking into account the information at hand must then be used to perform the record linkage. In this article, the outcome of using different rules for linking data from administrative registers into the statistical farm register is discussed.

Keywords: Record linkage, adminstrative registers

Title of abstract: Data integration to quantify occupational injury risk wage premium

Pawel Strawinski, Dorota Celińska
University of Warsaw, Poland

Studies of occupational injury risk influence on wages and wage differentiation are rare due to methodological and data problems. In this study we show how to take advantage of the surveys conducted within Polish Statistical Surveys Programme while analysing occupational risk. We combine information from two public data sources: the Structure of Earnings Survey (SES) and the Accidents at Work Data regained from the Statistical Cards on accident at work to estimate the influence of occupational injury risk premium on wages and the gender wage gap. While the SES is a well-known database among economists analysing labour market, the latter is a less popular one: it contains annual data about all of the accidents and their characteristics. Merging of the databases enables us to operate on the individual level instead of aggregated data. We match every observation from the SES data with estimated accident rates calculated separately for every economic section/occupation/gender cell. We show that it is feasible to work at three digit level of the International Standard Classification of Occupation and still obtain precise estimates of occupational injury risk for both genders. Contrary to the existing research, we take directly into account the differences in work related risk among workers. As an illustration, we use the combined data in occupational injury risk wage premium estimation, and answer the question to what extent the occupational injury risks explain the existing gender wage gaps.

Keywords: accident, occupational risk, compensating wage differentials, gender wage gap

Title of abstract: Data integration – idea for reduction of complexity

Natasa Saranovic, Saska Zivaljevic, Bojana Radevic
Statistical Office of Montenegro

MONSTAT conducts hundreds of surveys every year and prepares dissemination covering virtually every aspect of statistics in the Montenegro. Our mission is to provide timely, accurate, and useful statistics in service to the government of the Montenegro. Despite complexity of surveys and the large volume of data processed daily, the most important goal is to provide precise, timely and reliable data. The most usual way of processing data of a statistical survey is a, so-called, stovepipe principle, where the complete process is performed – from data entry to publishing the results separately for each survey. Therefore, NSIs have a situation where the processing of statistical data is done on various platforms, implemented on a variety of software tools, with data stored in various ways, separately for every survey, and even worse, for every statistical phase. Many of NSIs overspend on technology in the quest for getting value for the institution. They need to manage large IT structures, which are very expensive and slow, in order to approve capital investment and operational costs. One of the key distinguishing factors of statistical offices is data integration. Other organizations and enterprises may have better insight in individual sources but NSIs know best how to combine data and turn them into reliable figures describing the whole society. We gradually put all data on one platform. In that way we got all the functionalities of a warehouse. We made a uniform approach to all surveys, put all information about every piece of data in one place and made a single tool to handle all that information and data. That is IST Integrated system of data processing. IST was designed, developed and implemented in four NSIs and present example of international collaboration in data integration area.

Keywords: Data processing, data integration, IST

SPEED TALK SESSION 04 – QUALITY IN STATISTICAL DOMAINS

Title of abstract: Quality assessment and quality measurement for profiling in Germany

Roland Sturm
Federal Statistical Office, Germany (Destatis)

Official statistics in Germany is introducing profiling of enterprises as a new task of the statistical business register. Profiling in Germany aims to analyze the structure of enterprise groups in order to identify within enterprise groups the enterprises as defined by European law (Regulation 696/93). The staff for manual profiling is located in 14 Statistical Offices of the Laender. The tasks of the Federal Statistical Office are to coordinate the work, to evaluate and ensure good quality, to improve the methodology of profiling. The presentation will propose ways to deal with issues as:

Target population: Manual profiling requires high resources. Accordingly profilers should prioritize on "the most relevant" enterprise groups, expecting to find out about "the most relevant" enterprises. What are suited indicators for selecting profiling cases? How to assess the effects of profiling?

Workflow: How can 14 profiling institutions be coordinated in order to produce comparable results? How to assess effectiveness and efficiency of the methodology and the working procedures?

User orientation: Profiling is a means in order to provide statistical users in the statistical offices with appropriate statistical units, namely enterprises. How to identify the main points of interest of the users? How to store and provide the results of profiling to the users? How to integrate the user feedback in the adaption of the profiling methodology?

Structure and content: Profiling detects the enterprise perimeter – domain statisticians need a set of variables. How to organize the work share between profiling (detection of structures) and data collection (detection of variables)?

Keywords: profiling, enterprise, business register

Title of abstract: Survey quality: Response bias in retrospective questions

Mathias Revold, Mari With
Statistics Norway

Studies show that survey respondents' failure to remember events accurately (forgetting), as well as psychological mechanisms leading people to present themselves in a socially desirable manner (social desirability bias) leads to inaccurate results. We will look at self-reported electoral participation as a case to examine the impact of forgetting on social desirability bias, and thereby the quality of survey data. In 2014 and 2017 the Norwegian survey on living conditions included questions on participation in the last parliamentary election. As there were no national elections between the surveys both referred to the 2013-election. Thus, we asked about the same event shortly and several years after it happened. Self-reported electoral participation is higher than actual participation. Some studies show that socially desirable answers increase

when respondents forget events as time passes (Belli, Traugott, and Beckmann 2001). At the same time, the likelihood of forgetting decreases when events are regular and seen as important (Tourangeau, Rips, and Rasinski 2000). This could apply to electoral participation. We will look at differences in reported electoral participation between the surveys, overall and for demographic groups. These results can be compared to the actual turnout in the same groups by using the administrative register on electoral participation. Furthermore, the survey has a panel design which makes it possible to study changes in individuals' responses. Thus, we will study the impact of forgetting and social desirability bias when surveys refer to events years after they happened. If forgetting leads people to give more socially desirable answers, self-reported electoral participation should increase over time. However, previous studies have examined the impact over relatively short time periods. If respondents feel less social desirability pressure years after the event, the reported participation rate may decrease and be closer to the actual election results.

Keywords: survey methodology, social desirability bias, forgetting, self-reported electoral participation

Title of abstract: Machine Learning Techniques to Forecast Population Using Eurostat Data: An Exploratory Study

Álvaro Gómez Losada, Néstor Duch Brown
Joint Research Centre – European Commission, Spain

Machine learning (ML) is concerned with the algorithms that transform data into useful intelligence. ML algorithms play a key role in the big data era, making it possible to analyze vast amounts of heterogeneous data in order to extract unknown and potentially useful information and to help inform intelligent decisions. National Statistical Institutes provide high quality statistical information that has been gathered and processed based on international standards and appropriate data analysis procedures. However, this rich information may not be suited to analysis by traditional ML algorithms because the sizes of the typical datasets fall well below big data thresholds. In this study we explore the usefulness of ML techniques to forecast the 2016 Spanish population size by age class, using demographic data for the years from 2004 to 2015, obtained from Eurostat's web page. The following variables were used to perform forecasting: immigration, emigration, fertility, deaths, first-time marriage and life expectancy by years. This approach was compared with ARIMA and exponential smoothing (ETS) projections, as well as with a baseline landmark, which considered the 2015 population by age as the projection for 2016. The Mean Average Error (MAE) and Root Squared Mean Error (RMSE) quality measures were used to evaluate the accuracy of the estimated population pyramid. In all cases, the ML algorithm outperformed the other approaches. Although these results seem to be very promising, a few methodological issues should be discussed, such as the implementation of ML methodologies on small data sets, and the type and nature of variables used to infer the predicted outcome.

Keywords: Machine learning, forecasting, population pyramid, Eurostat, open data

Title of abstract: Survey on management practices

Paula Bordelo, Carlos Coimbra, Almiro Moreira, Sofia Rodrigues, André Sousa, Cristina Neves
Statistics Portugal

Statistics Portugal carried out an unprecedented survey over a sample of enterprises, established in the legal form of companies, achieving nearly four thousand valid replies, which made it possible to obtain information on management practices and characteristics in 2016. The survey, of a qualitative nature, falls within the scope of a range of statistical operations intended to disclose information on factors that, although with no explicit monetary reflection on enterprise accounting, constrain their competitiveness in a context of growing integration within overall economy. The main results obtained were divided according to four strata variables: Age of the enterprise, Belonging to an economic group, Size of the enterprise and Economic activity. Some examples of the main findings were:

- In 61.0% of the companies, top managers had a bachelor's or higher degree. This percentage was 82.9% in large enterprises and 43.7% in microenterprises.
- In about 70% of the companies, the top manager exercised functions under exclusivity (60.6% in microenterprises and 78.4% in large enterprises).
- In about 51% of the companies no promotions were awarded to employees with management functions. This percentage decreased to 44.3% in the case of employees with no management functions.

In order to obtain a synthetic measure of management quality, an indicator has been built (gscore), based on an indicator designed for a similar survey of the US Census Bureau. Combining the information from this survey with information reported to Statistics Portugal in other statistical operations, the results obtained indicate a significant relationship between management quality and the economic performance of the enterprises.

Keywords: Survey on management practices, synthetic measure, relationship between management quality and economic perfomance

Title of abstract: Improving efficiency of the sample design with sensitivity analysis of the thresholds in the Finnish horticultural survey

Anna-Kaisa Jaakkonen, Johanna Laiho-Kauranne, Mika Kuoppa-Aho
Natural Resources Institute Finland, Finland

The horticultural statistics are collected for the European agricultural policies based on the regulations on permanent crops and crop statistics. In FInland the horticultural survey has been conducted annually as a total survey with a threshold on standard economic output (SO) of the horticultural enterprises. We investigate the impact of increasing the SO threshold both on the quality and coverage of the final estimates, and on the survey costs. With the sensitivity analysis, wedemonstrate how to optimise threshold on the standard economic output to balance the survey costs and the quality criteria of the survey defined in the EU regulation survey for permanent crops. We also present the method of deriving the standard economic output for the horticultural enterprises in the sampling frame. To reduce the survey costs, and to improve the quality of the survey data, the sampling design of the survey is reviewed in detail. The data collection of the horticultural survey uses mixed-mode approach; using register data, web-survey, telephone interviews and accepts also paper questionnaires. With the increase of the thresholds we can also analyse the increase in the web-survey response rate. Thus it is expected that larger horticultural enterprises tend to respond through the web-survey more likely; while those who are interviewed tend to be on average smaller enterprises. Therefore, we will also present the impact of the efficient sample design on the expected improvement on the timeliness of the survey data. The improvement of the sampling designs is increasingly topical as there are extensively new information needs, and the statistical offices must balance between the statistical and response burden. Improving the efficiency of the sampling design is directly reducing the survey costs and the response burden.

Keywords: Sampling design, reduction of burden, survey costs, web surveys, survey thresholds

Title of abstract: Understanding the effect of the global economy on the Balance of Payment

Caroline Bo, Casper Winther, Anette Morgils Hertz
Statistics Denmark

Increasing our focus on large enterprises' production setups has been of great importance to heighten the quality of measuring global activities affecting the Danish economy. Two global production setups – merchanting and processing - are central. During 2014 when the BPM6 was being implemented, we found that there was significant underreporting of these types of activities. Hence, we were not catching the full scale of global activities by Danish enterprises affecting the Danish economy. We decided to start validating data on the largest cases, i.e. the Danish enterprises, expected to have most global production. By validating data on the largest companies on an ongoing basis, we now believe to have a trustworthy indication of how our large enterprises affect our Balance of payment. In this paper we will describe how we came to validate merchanting and processing data etc. in a validation across statistical domains as well as include a description of the practical approach to this work.

Keywords: Large enterprises, Large cases unit, Balance of payment, validation, BPM6

Title of abstract: 45-day flash estimates of a PEEI: the Italian job vacancy rate – methods, revisions, cyclical properties

Annalisa Lucarelli, Marina Sorrentino, Emilia Matera, Giuliano Latini, Diego Chianella
Italian National Institute of Statistics (ISTAT)

The EU regulation on quarterly job vacancy statistics requires data transmission within 70 and 45 days after the end of the reference quarter. The published indicator is the job vacancy rate, that is the ratio between the number of vacant posts and the sum of vacant and occupied posts, which is included among the Euro Principal European Economic Indicators (PEEIs) and is considered a potential leading indicator of the business cycle. The Italian job vacancy data are based on two direct business surveys and an auxiliary administrative based source (for editing and imputation and calibration). The procedure used to produce the data for the 70 day deadline makes full use of the reference quarter data from all three sources. However, for the 45 day deadline fewer data are available and as a consequence a different procedure needed to be developed and implemented. In particular, administrative based data for previous quarters are used, as well as more limited sets of respondents to the two direct surveys. The results have proven so far very satisfactory. The revisions between job vacancy rate estimates for the 45 and 70 day deadlines are often zero, especially at the higher aggregation levels. This happens also if the rate numerator and denominator change significantly between the two estimates, due to the different sets of direct survey respondents and the different populations on which the calibration constraints are based. Furthermore, the flash estimates job vacancy rate generally show good cyclical properties. The flash estimates quality, however, can be negatively affected by intense and prolonged downturns and upturns, when the impact of the use of calibration constraints based on previous quarters rather than the reference one can be more relevant. Improvements in the procedure to account for this limit could be studied in the future.

Keywords: flash estimates, vacancy rate, leading indicator, business survey, administrative data

SPEED TALK SESSION 05 – HUMAN RESOURCES IN A CHANGING STATISTICAL ENVIRONMENT; COMMUNICATING QUALITY OF OFFICIAL STATISTICS

Title of abstract: Statistics regulation for an age of "post-truth": The UK's new Office for Statistics Regulation

Ed Humpherson
UK Office for Statistics Regulation

The last 18 months have seen significant shifts in the UK's institutional environment, resulting from the Bean Review of Economic Statistics; the increased scrutiny placed on statistics in a shifting political context; and a broader society-level concern about a "post-truth" era. Into this environment, the UK Statistics Authority launched a new regulator – the Office for Statistics Regulation – in November 2016. This new body stands up for statistics in public; takes a whole-system view of statistics in areas like health, migration and crime; and raises the quality of individual statistical outputs. In this presentation, Ed Humpherson, the head of the Office for Statistics Regulation, will explain the vision for a public champion of statistics – to celebrate when statistics serve the public good and challenge publicly when they do not. He will explain what the Office for Statistics Regulation is doing to drive up trustworthiness, quality and value of whole systems of statistics; and how it has intervened publicly in high profile uses of statistics, including crime data, migration and the UK's contribution to the European Union. It's a story of organizational development; of reimagining statistics for a data-rich world; and of why statistical producers should have nothing to fear from the term "post truth".

Keywords: governance, regulation, post-truth, trustworthiness

Title of abstract: Statistics storytelling – A magical mirror: Israelis, Palestinians, Jordanians and Egyptians cooperate in order to improve their official Foreign Trade Statistics by using „Mirror Exercises"

Nili Karsai - Bibi, Helen Brusilovsky
Israel Central Bureau of Statistics (ICBS)

Official foreign trade statisticians deal with asymmetry problems in bilateral trade statistics. In an ideal situation, country A's data on exports to country B should be equal to country B's data on imports from country A. In real life there are discrepancies which create confusion among users of trade statistics. Such discrepancies should be reduced as much as possible since accurate and comparable trade statistics helps decision makers and facilitate actual trade. „Mirror Exercises" are a quality methodological tool used to identify reasons of such discrepancies and conduct reconciliations in order to explain and reduce asymmetries. The article reviews some basic concepts of official foreign trade in goods statistics needed to understand the subject. The issue is explained by a case example: Mirror exercises on Israeli, Palestinian, Jordanian and Egyptian official bilateral trade in goods statistics. The exercises were organized and supported by the European Union (EU) Euro-Mediterranean Statistical Cooperation Project (MEDSTAT). The main reasons for the discrepancies found in these mirror exercises were: The use of different trade systems, diverse data sources, misclassifications, data recording issues and errors. Quality improvements achieved in the exercises were as

follows, better implementation of international standards, improved coverage and greater harmonization of trade in goods statistics. Results of these exercises could be seen especially at fuel products data, where a major asymmetry was solved by analyzing data from trade documentation and sources. As a general advice, users of bilateral trade data are encouraged to check both countries official data. The paper was written in a user-friendly „Statistics Storytelling" style, as recommended by the United Nations (UN) guide for increasing literacy of official statistics among users. This aim is in accordance with the European Statistical System (ESS) vision to address user needs.

Keywords: Statistics storytelling, Bilateral trade statistics, Mirror exercises, Data asymmetries, Reconciliation study

Title of abstract: Quality quest in official statistics through international cooperation: Turkish Statistical Institute Experience

Gürdal Gedik
Turkish Statistical Institute

Quality is one of the most important phenomena in every field of our lives in this age. Since information and knowledge are fundamental pillars for modern societies and official statistics provide an information infrastructure for the needs of various categories of users, enhancing the quality of official statistics is of utmost importance in this frame as well. Accordingly, National Statistical Institutes (NSIs) have been intensively working to improve both the quality of their systems and products through various ways. However, high quality statistical information is neither easy nor cheap. NSIs need to build sound infrastructures and invest in the production of high quality statistics in order to do so. Budgetary constraints in the public sector have led to shrinking resources also for statistical administrations. These issues cause considerable challenges for official statistics. International cooperation in the field of statistics offers an important contribution to overcome these difficulties. In line with this perspective, the principle on coordination and cooperation was very recently included in the European Statistics Code of Practise as „Principle 0". Turkish Statistical Institute (TurkStat) has also been an active actor in the field of statistical cooperation. Having a long history in statistics and gained significant experience through European Union harmonisation process, TurkStat shares her knowledge with other NSIs in order to contribute to their quality of human resources as well as quality standards of official statistics, in a demand driven manner. This paper presents the international cooperation frameworks and practical implementations of TurkStat in the field of statistics. It explains how TurkStat supports the capacity building efforts in human resources and accordingly how it contributes to the improvement of the quality of statistical production in certain NSIs.

Keywords: Quality, international cooperation, human resources, international organisations, donor partners

Title of abstract: E-learning platform and innovative teaching methods in ISTAT

Tiziana Carrino, Patrizio Di Nicola
Italian National Institute of Statistics (ISTAT)

In the last years Istat invested in experimentation of innovative teaching methods and tools, aiming at building up learning environments based on the integration of "traditional" tools with Technology Enhanced Learning resources. Within this context, an e-learning platform was set up in order to promote the statistical culture and innovative teaching methods and to develop and enhance statistical skills. E-learning and blended-learning courses are characterized by an innovative instructional design. Videotutorials, videolessons, webinars are some of new teaching formats tested recently to enrich our training programs. These new innovative teaching methods and tools: support Istat modernization, promoting and facilitating the change; favor the statistics dissemination expecially in the National statistical system; improve training to the data collectors, allocated to all the Italian territory. E-learning methods are clearly connected with the competence system and the projects involved in it as 360 feedback, survey on occupation, etc. The paper will describe the e-learning platform focusing on the main methods and tools used in Istat. Paper will analyze also main advantages and critical issues of these new training methods.

Keywords: e-learning, e-learning platform, support to modernization

Title of abstract: Setting up a competence system at Istat

Pietro Scalisi, Patrizio Di Nicola, Antonio Ottaiano
Italian National Institute of Statistics (ISTAT)

Last year, the Italian Institute of Statistics implemented a new statistical production system and, accordingly, a new organizational model. Within the framework of this modernization process, new roles and new competences are required to achieve the Institute's goals. The development of human resources has to go along with the modernization of the statistical production and play an active role in corporate strategy and planning. To support this change, the Istat Human Resources Department is engaged in the definition of a competence system. This will help to better manage the human capital, foster competence development and create an organized source of information to support HR policies on recruiting, mobility, training planning, career development, etc. Istat competence system will be based on information gathered by means of various initiatives such as: survey on occupations - identification of the occupations carried out by staff; development of a technical competence data base – mapping of the knowledge and the technical skills possessed by staff and their level; survey on organizational skills – identification of the organizational competences used in working processes; identification of the expected competences – identification of the competences required in the working processes and measurement of gaps. The paper will describe the achievements and way forward to build a competence system and focus on strong and weak points.

Keywords: modernization, competences, human resources policies

Title of abstract: Revision Policy and Revision Analysis in Short Term Statistics in Istat: Recent Developments in Standardization and Communication Innovations

Maria Giovanna Piras, Donatella Tuzi
Italian National Institute of Statistics (ISTAT)

During the last two years the Italian National Institute of Statistics has deeply worked in finding out and establishing a common framework in revision policy and revision analysis in the short term statistics domain. The objective was the setting of an extensively shared tool for internal quality governance and the building of an easily accessible platform where information on practice and measure of revisions can be drawn by stakeholders. At this aim, a mandate was given to a transversal working group whose members were chosen from different structures of the Italian NSI: methodologists, statisticians, experts in IT and in communication, quality managers. Main result of this work, the setting of a dedicated section of Istat website where, beside the presentation of the established protocol on policy, analysis and dissemination of revisions, different tools and related measures of revisions based on "revision triangles" are provided. At the moment this new environment has completely been settled for three short terms production lines: Monthly Industrial Production, Quarterly National Accounts, Quarterly Labour Costs statistics. In the next future, it will be extended to cover all the short term themes, implying the overcoming of heterogeneous and extremely fragmented practices, towards the improvement of a wider transparency and accountability of the produced statistics. The purpose of this paper is threefold: presenting the main established protocol on revision policy, analysis and dissemination of revisions of published statistics; describing main input/output of a generalized tool for storing indicators and calculating revision indicators; identifying communication tools (calendar of revisions, release of vintage series, dissemination of the revision triangles, etc.).

Keywords: Revision Policy, Quality Indicators, Revision Triangles, Standardization, Transparency

Title of abstract: Engaging with users to improve quality

Catherine Bremner
Office for National Statistics, United Kingdom

'Quality is everyone's responsibility.' Do you review processes to ensure the best quality data is produced? Regardless of the topic, speaking to our users to improve the quality of a process or output is an important step. There is a requirement under the UK Code of Practice for Official Statistics to: 'Report annually the estimated costs (for example, on businesses, service providers, or the public) of responding to statistical surveys and strive to develop methods that will reduce the costs to individual organisations or people.' The ONS collects compliance information from all Government Departments in the UK (excluding Scotland) and publishes it in a report called the Online List of Government Statistical Surveys. This publication was set up a number of years ago to meet the Code of Practice obligation. Like any output the quality should be assessed on a regular basis. Current metrics on the use of the Online List are minimal; this is an indication that the quality needs to be improved. A list of all Government Statistical surveys and their compliance is beneficial for a multitude of reasons, therefore, a user review needed to be conducted to improve quality, exposure and capability. The processes are being reviewed by Methodology to ensure the current methodology is fit for purpose, the publication itself is having a facelift, and users are being asked what they want.

The main focus of the user review is to explore:

- who uses compliance information,
- what is compliance information used for,
- is there any additional information users would like.

The paper details users and data compilers feedback, along with recommendations made to improve the publication and prevent misuse of data.

Keywords: user review, users, quality reviews, compliance

SPEED TALK SESSION 06 – QUALITY & INNOVATIVE SOURCES AND METHODS IN STATISTICS

Title of abstract: Mobility survey on metropolitan areas - innovation on methods and procedures

Cristina Neves
Statistics Portugal

Statistics Portugal implemented a Mobility survey on metropolitan areas, aiming to get data concerning urban trips' patterns of the population according to residence area and motivations. In the particular case of Metropolitan Areas (MA), where the complementarity of transport modes is essential especially for commuters, the mobility surveys are important tools to support the decision in the transport system scope, namely in what concerns the intermodal network definition and ticketing systems. The survey was defined in order to produce results to feed the national and European statistical systems, as official statistics, and to obtain information in the most possible harmonized way according to EU recommendations for mobility statistics. Given the specificities of the MA in terms of population and territory characteristics, an innovative approach in terms of methods and procedures have been taken into account by Statistics Portugal in the methodological study of the survey. A study on territorial division was conducted, in order to allow the definition of homogeneous accessibility areas, via a cluster analysis to group parishes with similar characteristics in terms of mobility, using the combination of Census data about the population and commuting and also the geographic data about transport infrastructure: rail and metro stations; main road intersections; inland waterways connections, on which buffers on their influence were created. Those homogeneous areas have been taken into account for the sampling design. Also an innovative approach in terms of sampling design was considered, in order to allow a combination of different data collection methods, namely computer assisted web interview (CAWI) and computer assisted personal interview (CAPI). A stratified two stage sample was implemented, with systematic sampling selection for the CAWI in a first phase and, in a second phase, a sub sample from the first phase was selected for CAPI.

Keywords: Mobility; clusters; buffers; accessibility; intermodal

Title of abstract: Digital transformation as an opportunity to become more relevant

Thorsten Tümmler, Thomas Riede
Federal Statistical Office of Germany

The digital transformation resounds throughout the land – it encompasses the economy, it pervades society as well as everyday life and it forces public administration, which generally lags behind, to stretch and catch up. Destatis wants to view the digital transformation not as a threat to its existing business model but as an opportunity to alter limitations of existing (legal) frame conditions and thus shape its future. Digital transformation however, is not an end in itself. It serves to better fulfill the strategic goals of Destatis, prime among them the quality and the relevance of statistical data. Therefore Destatis has developed a Digital Agenda that links the ambition for the digital transformation to the strategic goals. The Digital Agenda also provides a road map towards this ambition by outlining key measures that need to be undertaken in order

to successfully transform. The Digital Agenda shall be revised every year in order to adapt ambition and key measures to new requirements as well as to reflect progress achieved. This paper aims to present the Digital Agenda and to outline its effect on the quality and relevance of statistics.

Keywords: digital transformation, digitalisation, strategic goals, relevance

Title of abstract: Data collection quality control using paradata and geolocation

Marcos Rodrigues Pinto
BGE – Brazilian Institute of Geography and Statistics

This paper describes the progress made by the Brazilian Official Statistics Office (IBGE - Institute of Geography and Statistics) to assure quality on its latest census. Management information systems are being extensively improved to deliver critical information to all the levels of administration on the data collection phase, producing important data for decision making. Tools like enumerator tracking by geolocation brings new possibilities to field monitoring, as it makes available the enumeration area coverage, not seen solely with the household coordinates registered before an interview on a CAPI approach. Also, the use of paradata (data about the process of data collection, produced by the enumerator) has shown it´s extremely promising value on quality control and consequent cost saving. The behaviour of the enumerator on the use of the handheld device can give detailed information concerning quality and productiveness. Examples of data that are currently used in Brazil are the record of coordinates during questionnaire execution, enumerator navigation on the questionnaire, interview length and questions sequence. Analysing this data while the collection is in progress and sharing with the field supervision has been shown very important as they can act preventively, checking an enumerator work that is signed as suspicious (possibly with maid up questionnaires), and making correct decisions. This work shows the potential use of paradata and geolocation as central piece of management information systems to produce high quality official statistics.

Keywords: paradata, data collection, geolocation, management information systems

Title of abstract: Building the territorial statistical register: quality control on geocoded administrative data

Eduard Suñé Luís, Daniel Ibáñez Vidal, Mireia Farré Mallofré, Cristina Rovira Trepat, Mònica Gasulla Ramon, Frederic Udina i Abelló
Institut d'Estadística de Catalunya (Idescat)

The Statistical Institute of Catalonia (Idescat) is developing a statistical information system based on administrative registers. All these registers usually contain a set of variables defining the spatial location of the microdata through a postal address. One of the subsystems used is the information about real estate cadastre, which contains around 10 million georeferenced estates. Our aim is to analyse its statistical quality. A set of polygons was created, associated with the concave or convex hull from the centroids of the estates grouped together by roads, as a first step in establishing the quality of the positions. One of the tasks carried out is that of the detection of outliers in these polygons. We focus on analysing the distribution of the distances between the points of each polygon. We begin with the idea of creating an indicator to rank the polygons according to their level of reliability, indicating the possibility of there being an outlier in the polygon. Our first concern is to identify the polygons with the most extreme values on the indicator, both those which could contain suspicious points (around 5%) and those we are sure do not contain any (approximately 70%).

By means of a comparison using multivariate analysis and anomaly detection techniques, we aim to verify that the results are almost the same. Secondly, our concern is to study the cases found in the intermediate zone, and in particular to check in which threshold we will obtain the optimal relation between the well--classified and the badly-classified ones. In order to evaluate our indicator, we must select a sample of cases and study them manually to ascertain whether they are correct or not. It is therefore in our interest that the sample mainly contains polygons found in the intermediate zone of the indicator.

Keywords: Registration of territory, georeferencing, statistical quality, detection of outliers

Title of abstract: Does Big Data mean Big Problems, or Bigger Opportunities?

Pedro Cunha, Sónia Quaresma, Jorge Magalhães
Statistics Portugal

The traditional approach to design a survey for the production of a deliverable is no longer maintainable. Life has changed in the last years not only on the perception of how the use of administrative data could lessen the burden on the respondents and the survey costs but also as the gathering of information through sensors and other automated devices proved to be a cheap and useful resource, shortening the time to produce statistics thus increasing its timeliness. Many questions arise when we try to use data, be it administrative or sensor generated, to produce official statistics. Most of them because the data was not collected with statistical purposes and methodologically its use can be complex as it does not meet statistical standards on concepts or definitions. For this reason our approach to BigData has been cautious and gradual. After the census 2011 operation a national dwelling register was built and since then this database has been fed with administrative data. We are now considering its enrichment using electricity consumption information collected through smart meters. This BigData source could assist us in determining dwelling occupation. Due to several restrictions on the BigData source availability it is not possible to establish a connection per household but instead to validate the coverage of areas. The resulting impossibility to perform microdata linkage lead us to study different ways in which to improve quality on the register, validate under or over cover areas and to develop new methodologies for sources combination. Different quality checks to the data extracted from BigData sources have to be applied and new combining methods are in this paper addressed in the specific context of the national dwelling register. Through the presentation of this case study we share the obstacles and gold nuggets found along the road of BigData exploration.

Keywords: National Dwelling Database; Bid Data; Partnership between Public and Private Organizations regarding Information collection; Smart Meters; Register Quality; Data integration

Title of abstract: Local decisions and new guidelines of the Official Statistics. A pilot study on Early School Leavers to explore the quality of integrated administrative data used to support local policies

Daniela Ferrazza, Simona Ballabio, Sara Casacci, Arianna Carra, Flavio Verrecchia, Alberto Vitalini, Lorena Viviano
Italian National Institute of Statistics (ISTAT)

The work is based on the experimental use of data from the ArchIMEDe project (ARCHivio Integrato di Microdati Economici e Demografici – i.e. Integrated Archive of Economic and Demographic Microdata) of the Italian Institute of Statistics (ISTAT), created with the aim of expanding the supply of information of regional and local interest through the production of micro-data from administrative sources. In a scenario of possible integration and/or transition from survey sources to administrative sources of Italian official statistics, the first step in the study of phenomena is the problematization of operational measures and existing definitions. Although the data used - coming from secondary sources - are not structurally similar to those of ISTAT sample surveys and are not even directly comparable, the new boundaries of official statistics show an increasing interrelationship which, at local level, requires an integrated reading of phenomena through the different sources available in order to obtain the best information and the best quality. The contribution, conceived as part of the SPoT project (data and methodologies for the development of Statistics for Policies on the Territory), proposes a case study of Early School Leavers, defined as young people between 18 and 24 years of age who do not have a secondary school qualification and who are not enrolled in an education cycle. The paper presents the results of some analyses carried out on residents in Lombardy using spatial association measures and hot spot maps. These results highlight, at the same time, elements which are useful from a substantial point of view for the understanding of the phenomenon at a local level, as well as methodological aspects relating to strengths and weaknesses in experimental use of secondary sources in statistics to support local policies.

Keywords: Experimental statistics, administrative data, ArchIMEDe, Early School Leavers, sub-regional and local data

Title of abstract: EBALUA: Building a new application to assist Quality Assessment of Statistics in Eustat

Enrique Moran Alaez, Maria Victoria Garcia Olea
Basque Statistics Institute – EUSTAT

"Ebalua", the Basque word for "Assess", is the provisional brand name of a new application that the Basque Statistics Institute-Eustat is currently building in order to assist the statistical staff to assess the quality of statistics produced inside the Basque Statistics Organization. Assessment is seen in Eustat as an internal reflection that the responsible for the statistics makes with his/her team about the quality of them so that some improvements for the new cycle can be introduced. The application provides the entire set of tools that the statistics team need to conduct the assessment, the questionnaires to fill in -based on the DESAP formdeveloped by Statistics of Lithuania-, the assessment report, the access to the data base with previous assessments,..., following the assessment protocol adopted by Eustat. The application is developed on the Internet because it is open to the Basque Statistical Organization, which includes Eustat and the statistical staff from the Basque Government Departments, in case this staff is part of a statistics team under Eustat's responsibility. EBALUA is integrated in the information system of Eustat so that it can take several programming data from it: the names of the statistics and of the team members, e-mail addresses, planned dates of the assessment, and so on. EBALUA communicates the start and the end of the assessment process to the statistics team and other people, such as the heads of the departments involved and Eustat's managers, sending messages and the assessments reports. EBALUA will store all the information in an Oracle data base and produce statistical reports of the assessment process, duration, delays an main features, giving a precise idea of the spreading of the culture of quality assessment in Eustat.

Keywords: Quality assessment, statistics team, assessment application, improvement, integration

Title of abstract: Improving official statistics with credit and debit card data

María Valdivieso Anívarro, María Rosalía Vicente
University of Oviedo, Spain

This paper explores the use of credit and debit card transactions to measure consumption and poverty, the latter understood as the absence of consumption or the restriction of it. These data are available on a timely basis and provide detailed information about spending activity with a finer granularity than that available in traditional household surveys. Accordingly, the exploitation of these data source can provide novel insights about both consumption patterns and poverty. Specifically, we estimate spending activity by the smallest geographical unit possible and investigate differences by age and gender. We also analyze the spatial distribution of these variables and how they relate to the macroeconomic features of each district. Finally, our estimates are compared with the official figures from household surveys.

SPEED TALK SESSION 07 – INNOVATION AND QUALITY IMPROVEMENT IN STATISTICAL PRODUCTION

Title of abstract: Quality monitoring and improvement of innovation survey in Republic of Macedonia

Tatjana Drangovska
State Statistical Office, Skopje, Macedonia

Innovation has a central role in the process of output creation and productivity growth. Data on innovation are widely used by policy makers for policy creation and by researchers for estimating of different aspects of innovation process. Because of that production of data on innovation with good quality is of a high interest. Starting from 2013, SSO started with conduction on survey on innovation. In order to ensure international comparability, the standard CIS questionnaire and Oslo Manual are used. For quality monitoring of the outputs of innovation data and also for making the figures transparent and easily accessible to the users, ESQRS, DESAP self-assessment and national Quality Report are prepared. The purpose of the paper is to present the improvement of the innovation data through the period based on the kay indicators calculated in Quality reports used as input for the future work. At the same time in the paper will be briefly presented the methods undertaken for improving the quality of output data for innovation, for example cross-checking method with data from other statistical surveys and the usage of data from administrative sources.

Keywords: Innovation survey, quality reports, quality improvements

Title of abstract: A territorial model for the management of data collection processes: a case study

Loredana De Gaetano, Pasquale Papa, Angela Maria Digrandi
Italian National Institute of Statistics (ISTAT)

The Italian National Statistical Institute (ISTAT) carries out the survey on the maritime transport of goods and passengers, as required by the Regulation (EU) No 1090/2010 of the European Parliament and of the Council. The survey is a census as it refers to the final overall amount of the arrivals and departures recorded in Italian ports. It states, according to law, the obligation for all the agents, consignees and shippers operating in Italian ports to provide to ISTAT the main information on the travels of the relevant ships as well as on the quantity of goods and passengers transported. In order to improve the quality of data collection, within the new ISTAT organizational structure, introduced during 2016, was created the Central Directorate for data collection (DCRD) specifically dedicated to the design, organization, management and integration of the data collection activities. Then a specific project assigned to a selected number of ISTAT territorial offices the new role, of conducting data collection activities on the territory. For the specific purposes of maritime transport survey, ISTAT territorial offices are entrusted with the tasks of updating the list of respondents and carrying out and monitoring data collection on the territory. In addition, according to this model a specific territorial office acquires the role to coordinate the conduction of data collection activity of all the other territorial offices. The paper presents the case study of the 'Campania' territorial office that, starting from the survey editions 2017 and 2018, co-ordinates the activities of data collection, in collaboration with

the ISTAT Service for the management of data collection from direct surveys (RDC). The analysis is aimed at presenting the improvements both in terms of process efficiency and quality of the results (response rates and timeliness) expected from the new data collection approach.

Keywords: Data collection, maritime transport, response rates, process efficiency

Title of abstract: Login on Smartphones: A triviality?

Johannes Volk, Karen Blanke
Federal Statistical Office of Germany (Destatis)

There is an ongoing discussion about transforming questionnaires into an appropriate mobile version in order to avoid mode-effects. However, a crucial point is frequently neglected or less discussed implementing online questionnaires: People have problems to get access to the online instrument, due to an uncomfortable login-process. So far, it seems to be rather a blackbox, why people aren't successful, since basic recommendations are applied (e.g. number of characters, which symbols to avoid etc.). Based on qualitative pretesting research shows concrete evidence on the design of an easy-done login on a Smartphone. By four waves iterative testing the login-process has been adjusted. Results touch topics like cryptic vs. catchy password, displayed vs. faded out input of password, proportional vs. non proportional font. Moreover, we analyzed the duration of the login-process, the key orders on different key boards, the design of input fields, the layout of the covering letter concerning login-information and test-persons' reported subjective effort and security-ratings. In the end we developed a combination of login design elements, we would prefer. We identified a lot of different details which can be addressed to simplify login processes systematically. In conclusion Charles Eames' word is even true for the design of login-processes: "The details are not the details. They make the design."

Keywords: Smartphone-design, online data collection tools; login-process

Title of abstract: About the Quality of the 2013 Census of Population, Households and Dwellings in Bosnia and Herzegovina: Statistical Results versus Political Controversies

Edin Šabanović
Agency for Statistics of Bosnia and Herzegovina

Census of Population in Bosnia and Herzegovina was conducted in 2013. Statistical offices conducted census in line with international standards. However, almost all main issues of the Census were politicized. The Census was given many non-statistical characteristics. It was presented by various interest groups as politically important activity, as a population register or a source of data, which could be used for tax policy or property management. For all these reasons, census activities were a subject of political controversies, topics for evaluation of several, mostly statistically unprofessional interest groups or individuals, who had more media attention than statistical expertise and whose influence to the attitude of Bosnian citizens was non--ignorable. The Post-enumeration survey (PES) was conducted in November 2013. The objective of the PES is to provide indicators of the census. In this paper, we will show how the PES was designed and conducted and which statistical models and methods were used for linkage of Census and PES data and for the analysis of Census quality. The last part of the paper will discuss the acceptance of the Census results by different

data users such as statistical institutions, journalists, general and expert public. This part of the paper aims to give a rough picture of the general statistical culture in Bosnia and Herzegovina through the prism of the Population Census. In conclusion the evaluation of the quality of 2013 Census of Population, Households and Dwellings in Bosnia and Herzegovina will be done as well as some proposals for future work.

Keywords: census, quality, coverage, content

Title of abstract: Flash estimates and increasing data quality as a result of modern IT-applications and job satisfaction

Sabine Schuster, Thomas Karner
Statistics Austria

From 2017 onwards, Statistics Austria publishes flash estimates for road freight transport statistics. These flash estimates enable data users to get first key figures of freight transport of Austrian road freight vehicles already one month after the reference quarter (t+1) instead of five months (t+5). Road freight transport statistics is – with reference to Council Regulation (EC) No 70/2012 on statistical returns in respect of the carriage of goods by road (recast) – performed as a sample survey. Weekly, questionnaires are sent to the respondents, who have to report all journeys of the selected trucks and return the web- or paper-based form to Statistics Austria two weeks later. To check the plausibility of the filled-in forms a Java-application is used since 2014. Prior to that, the reprocessing was very complicated and included different steps and different people in charge. Now, each questionnaire is attributed to one staff member, who is registering its completeness, checking its plausibility and - if necessary - is getting in contact with the respondent to clarify certain parameters. Once finalised, each questionnaire is directly added to the authentic database. Following the two-factor theory of Herzberg about job satisfaction, it was noticed, that the motivation of the staff members improved due to intrinsic factors such as increased responsibility and decision autonomy. The work tasks have become more demanding and therefore interesting with a better insight and understanding of the whole process. Thereby the quality and availability of the produced data enhanced. It was analysed, that one month after the reference quarter already more than 70% of the questionnaires could be correctly finalised. Calculated flash estimates showed similar outcomes to the final key results with a slightly higher sampling error which is still in the frame of precision requirements.

Keywords: Motivation, Quality, Flash Estimates, IT-Applications, Road Freight Transport Statistics

Title of abstract: Do declining response rates negatively affect sample composition? A longitudinal analysis using data from the German General Social Survey (ALLBUS)

Michael Blohm, Achim Koch
GESIS - Leibniz Institute for the Social Sciences, Mannheim, Germany

The German General Social Survey (ALLBUS) is a repeated cross-sectional multi–thematic survey conducted every other year since 1980. Like many other surveys ALLBUS has been facing the problem of declining response rates. Since 1994 ALLBUS response rates decreased steadily from 54% to 35% at present – despite exerting higher fieldwork efforts. Declining response rates give rise to the question whether this affects survey data quality. Selective participation behavior might threaten the representativeness of a survey in general or with regard to specific questions. In our presentation we use ALLBUS data to analyze the rela-

tionship between the response rate and indicators of nonresponse bias over time. The ALLBUS provides a good basis to analyze this relationship, since it has used an almost identical study design since 1994. This includes the definition of the target population (adults living in private households in Germany), the sample design (samples of named individuals), the net sample size (3.500 completed interviews), the interview duration (around 70 minutes), and – particularly important in this context – a consistent calculation of response rates. To investigate nonresponse bias we analyse a) survey participation in ALLBUS with sample frame information from the population registers and b) deviations of the net sample from external benchmark data (namely the German (micro-) census from official statistics). First results indicate that decreasing response rates are not systematically associated with deterioration in demographic sample composition. We will argue that these findings only hold in "controlled" sampling design settings (for instance, when samples of named individuals have been used).

Keywords: Response rate, Nonresponse bias, ALLBUS

Title of abstract: Maintaining good response rates while changing survey techniques: the case of the Italian Quarterly Business Survey on Job Vacancies and Hours Worked

Marina Sorrentino, Emilia Matera, Latini Giuliano, Diego Chianella, Annalisa Lucarelli
Italian National Institute of Statistics (ISTAT)

Response rates in Italian business surveys range from 30-40% in structural more complex ones to a maximum of around 85% in few monthly panel surveys involving only large enterprises or carried out so far exclusively via CATI (Computer Assisted Telephone Interview). The Italian Quarterly Business Survey on Job Vacancies and Hours Worked had always had response rates in the higher range of this interval (often around 65-70%). This was achieved through a mixed mode survey technique (CATI for around 70% of respondents and CAWI, or Computer Assisted Web Interview, for the remaining 30%), carefully timed and addressed email reminders and daily interaction with the sample enterprises by survey experts. Changes in the laws concerning the interactions between enterprises and public administration and financial constraints made it impossible to continue using the CATI technique after 2016 and required the switch to a unique response mode, CAWI. This change posed serious risks of a fall in response rates. A series of measures were taken to avoid this outcome. In particular, email and certified email reminders were intensified, increased resources were dedicated to responses to emails, certified emails and phone calls by sample enterprises so as to improve their already good timeliness and accuracy. Furthermore, starting from the survey for the II quarter 2017, telephone reminders carried out by an external provider have been introduced, where the telephone operators can also support the enterprises on the questionnaires requests, access to the relevant website and filling in of the online form. The results in terms of response rates have so far been encouraging.

Keywords: business survey, response rate, survey technique, CATI, CAWI

Title of abstract: Design and analysis of the improved Poisson and negative binomial item count techniques

Barbara Kowalczyk – SGH Warsaw School of Economics, Poland
Robert Wieczorkowski – Statistics Poland

Reliable data on stigmatizing, socially unaccepted or illegal features are very hard to obtain in direct questioning. Many indirect methods of questioning have been developed to help in eliciting honest answers to sensitive questions and to eliminate the social desirability bias. Item count techniques (ICTs) pioneered by Miller (1984) constitute an example of indirect survey techniques designed to deal with sensitive features. These techniques have many practical advantages: they are easy to implement, can be used in telephone and internet surveys, and the way they protect privacy is easy to understand by respondents. Recently Tian et al. (2017) proposed new item count techniques called Poisson and negative binomial ICTs. The new methods give many opportunities for further theoretical and practical developments. But the methods require large sample sizes to obtain a reasonable precision. Efficiency is an important issue in indirect methods of questioning. Protection of respondents' privacy is usually achieved at the expense of the efficiency of the estimation. In the present paper we propose new improved Poisson and negative binomial ICTs in which each of the two subsamples serve both as a control and a treatment group. This procedure allows to increase efficiency of the estimation as compared to the classic Poisson and negative binomial ICTs and maintain respondents' privacy at the same level. In the paper we introduce methodology of the proposed improved method and accompanying statistical theory. We analyze and compare best linear unbiased and maximum likelihood estimators of the population proportion of the sensitive attribute. We also compare the improved technique with previously proposed classic Poisson and negative binomial ICTs. The improvement is obtained in terms of efficiency. Gain in efficiency is achieved without affecting the privacy of respondents. Theoretical results presented in the paper are illustrated by comprehensive simulation studies.

Keywords: sensitive questions, indirect questioning, item count techniques

SPEED TALK SESSION 08 – QUALITY OF MULTI-SOURCES STATISTICS

Title of abstract: Balance of Payments quality management in Poland

Pawel Michalik, Jacek Kocerka
Narodowy Bank Polski, Poland

The BoP quality management process starts with the source data that must be collected from reporters. Special focus will be given to clarity of reporting requirements, close contact with reporters, training provided to reporters and sanctions for non-compliance. The second stage consists of collecting data from the Central Statistical Office of Poland and administrative data from other institutions. Cooperation quality has a direct impact on quality of statistical process. Institutional bodies in Poland that facilitate exchange of knowledge and information will be described. The next stage ie.. compilation process will be presented with tools and techniques for detection of outliers and other non-standard observations. The last stage of the process, ie. data dissemination of statistics with modern tools and techniques dedicated to key user groups will be described. Special emphasis will be given to feedback received.

Keywords: Balance of Payments, quality menagment, central bank statistics

Title of abstract: Data validation in livestock statistics

Esa Katajamäki
Natural Resources Institute Finland

Natural Resources Institute Finland made a study on data validation of livestock statistics in 2017. The key goal of this project was to evaluate the validation methods that are currently in use and to make recommendations for improving them. The statistical process from administrative register to statistics and validation methods in different stages were developed during this project. Another objective was to evaluate and update revisions related to direct data collection. The livestock statistics validation methods currently in use were documented and their strengths and weaknesses were evaluated. In the comparison with other agricultural statistics, the validation methods for administrative register data were compared to methods used for agricultural structure statistics and horticultural statistics. For administrative materials, the project recommends implementing a unified process pursuant to the model which was developed in this project. Based on the model, register data is first copied on unit level into the raw data table of the statistics. The data is then transferred to the basic data table on unit level. All discovered errors and missing data is corrected in the basic data table. At the next stage, the publishable sum level data is calculated and transferred to the sum level database table. A material revision report and different authenticators and validation rules are connected with each stage. In statistics production that is carried out using this process the most important step in data validation is to ensure that the unit level data in the basic data table is correct. The sum level revisions performed after this are a last check that the data to be published is accurate. If data errors are discovered after publishing, the required corrections are made to the unit level basic data.

Keywords: Data validation, livestoc statistics, administrative register

Title of abstract: Impact of estimated and administrative data on quality

Zsolt Kovari
Hungarian Central Statistical Office

The growing interest in the service sector compels to using of model based methods. These procedures have impacts on the quality and it is an exciting question how these effects can be measured and analysed. There are general methods available to us but the application of these means always more or less new methodological challenges. Hungarian Central Statistical Office has service turnover date from 2009 using the same nomenclatures and classification systems. The new group was created in Hungarian Central Statistical Office in order to process the VAT database and produce reliable administrative datasets for further use. How can the quality of results are measured? The paper presents the challenges with which we faced over processing VAT data summarises and compares the tested methods and describes the main achievements, outlines a new methodology of the VAT data and introduce the main effects for the quality of the estimated data.

Keywords: Administrative data, VAT data, data cleaning

Title of abstract: Implementation of a General Statistical Business Process Model in the administrative resource

Elsa Dhuli
Institute of Statistics, Tirana, Albania

The quality of production of official statistics is a contemporary demand which is growing dramatically by all countries, no matter what their inputs, statistical surveys or administrative sources are. Increasing the use of administrative resources as a primary or substitute element first requires first-rate quality assessment. Given the fact that the common approach is the measurement of the input quality and the statistical product process, when the main source is the administrative source, this approach takes on other dimensions. Generally speaking, the quality of each processing phase up to the realization of the statistical product has its importance, but the quality of the collection phase of administrative information and the implementation of statistical models in their holders also significantly guarantees the quality of other phases of the process. This paper aims to outline the implementation of a General Statistical Business Process Model in the administrative resource system at some of the phases to guaranty their quality to be usable as primary source for producing official statistics. The analysis of the quality of each phase of the statistical process from the administrative source of the experience of developed countries and the recommendations given for further improvement is a positive approach for enlargement countries.

Keywords: Quality, administrative source, GSBPM

Title of abstract: Who's telling the truth? Statistical techniques for error detection in double-sided reporting of money market transactions

Gemma Agostoni, Sébastien Pérez-Duarte, Steven Cassimon
European Central Bank, Frankfurt am Main, Germany

In 2016 the European Central Bank started collecting statistical data on money market transactions based on the Money Market Statistical Reporting (MMSR) Regulation. This granular dataset covers four segments of the euro money markets, namely unsecured, secured, foreign exchange swaps, and euro overnight index swaps. The detailed trade data to be provided comprise amongst others the volume, rate, and counterparty or collateral type information together with the time when the transaction was conducted. On average 45,000 transactions are being received every day from the largest 52 banking institutions in the euro area. Different procedures to ensure high quality have been adopted, including using ISO 20022 XML standard for the exchange of information. Particularly relevant are the methods aiming to reconcile the data reported by the different institutions. Both sides of a transaction (borrowing and lending) are reported by the parties involved. The lack of a unique transaction identifier poses significant challenges to the identification of the two sides of a single transaction. This paper presents different techniques, applied to MMSR, for pairing and matching the two sides of a transaction based on incomplete and partially incorrect information. Errors can be of several kinds: over-reporting and under-reporting of transactions, and misreporting. First, the matching exercise is conducted at macro level comparing the total volume exchanged between each pair of counterparties. This allows identifying the institutions experiencing possible issues of under- or over-reporting of transactions. In a second step, absolute, partial, iterative, and fuzzy matching techniques for pairing and matching individual transactions are developed, which allow automatically identifying out-of-scope or missing transactions, and misreported values, which would have not been detected otherwise. Inconsistencies detected are then referred to the specific reporting agents. These techniques have been critical in enhancing the quality of the MMSR, and may find applications in other datasets.

Keywords: Transaction-level data, double-sided reporting, pairing and matching techniques, under- and over- reporting, misreporting

Title of abstract: Boomerang effect of quality control on the compilation of Financial Accounts and flow of funds – the experience of Banco de Portugal

António Agostinho, Susana Santos
Banco de Portugal, Lisboa

Financial Accounts are fundamental to monitor financial stability by quantifying the impact of financial decisions of a host of economic agents. In Portugal, the compilation of these statistics is a responsibility of Banco de Portugal. One of the main purposes of the Statistics Department of Banco de Portugal is to ensure this statistical production with high quality standards, aiming at fully meeting user's needs, by developing a wide set of quality control procedures. Financial accounts are derived statistics stemmed from a vast array of other primary statistics, including balance of payments and monetary and financial statistics. In this context, Banco de Portugal developed a multidisciplinary team with experts from financial accounts and from the different underlying primary statistics. Within this format, all team members are co-responsible for producing national financial accounts, on a bottom-up approach, thus improving both the quality of these statistics, as well as the quality of primary statistics. This is the result of a systematic iterative process of data cross-check and reconciliation which may represent an opportunity to validate the soundness of microdata, on a top-down approach. To better understanding economic sectors' interlinkages and to assess how intersectoral financial linkages have changed, flow of funds is a powerful analytical tool.

Keywords: quality control, financial accounts, multidisciplinary team, data cross-check, flow of funds

SPEED TALK SESSION 09 – QUALITY OF STATISTICAL PRODUCTION PROCESSES

Title of abstract: Calendar effects in a monthly-based CATI survey

Antonio J. Rueda Clausell
National Statistics Institute of Spain (INE)

CATI is the acronym of Computer Assisted Telephone Interviewing, a collecting method used in many surveys since it is cheaper than face-to-face methods, but it has the disadvantage that we do not know anything about the household or person we call until they are contacted. In any CATI survey, at the beginning of the collecting period, the sampling units have a probability of being contacted. This probability depends on many different factors, but in any case, it decreases as the period goes on and the most difficult units remain without contact. Since it is more difficult to contact these units, interviewers use less time to conduct interviews and have more time to make new telephone calls. As a consequence, the more difficult the contact is, the more calls can be done. This gives them the possibility to make more contacts attempts. We studied the evolution of the probability of contact in a monthly-based survey, and how it acts on the number of telephone calls. On the one hand, the number of respondents increases because they are contacted at some point, although some of them refuse when they notice the large number of missed calls they receive. On the other hand, interviewers are discouraged when they make a lot of unproductive calls. Knowing how the number of calls impacts on the response rate, we can size the call center to avoid the discouragement of the interviewers as well as increase the final effective sample. It was not the aim of this study to get the "optimal" number of interviewers throughout the collecting period, but only to do a descriptive analysis of the influence of this on the results of the survey.

Keywords: CATI, data collection, response rate

Title of abstract: The economy has changed, the official statistics also change: the new business statistics system

Alessandro Faramondi, Serena Migliardo, Giampiero Siesto, Roberto Sanzo, Emanuela Trinca, Roberto Di Manno, Stefano Menghinello
Italian National Institute of Statistics (ISTAT)

Changes that happened in the last decades to economic level have changed the economies of the single countries, included in a single globalized economic system. The globalization and the role of multinationals highlighted the fragility of the measuring systems of the main economic variables about official statistics. The system of legal units as analysis unit shows its limits. The economic process splitting of business groups, based on the structure of the legal units, highlights conceptual inconsistencies, measurement problems, linked to the duplications of flows and to the ancillary and vertically integrated units. The new system, implemented since 2019, requires the aggregation of the legal units of the same group, when they don't have decision autonomy. This will impact significantly on the production system measurement. There will be a reduction in the number of companies, an increase in the size of the same and the reduction of some economic aggregates, because of the consolidation of the intra-flows. There is also a reshaping of value added and productivity per activity economic. The new plant will have a lot of implications and will invest

all the areas that deal with economic statistics. In this perspective, the process foreseen a strong integration among administrative sources, business survey, business register and EGR. In this frame is strongly used a profiling approach. These changes will impact on territorial statistics. Starting from the national data, micro-level information must be obtained. The income approach is used to calculate value added. With this work we want to explain the whole process undertaken to define new estimation in the area of business statistics in Italy to implement new BS system.

Keywords: Business Statistics, Sources Integration, Profiling, Internal Production Process, Micro data Approach

Title of abstract: Improved quality in statistical production by using standard processes in phases: design, build and collect

Biljana Ristevska Karajovanovikj, Valentina Todorova Hristovska
State Statistical Office of the Republic of Macedonia (SSO)

Metadata describe comprehensively statistical data and processes, assist in their interpretation and they are fully integrated in statistical processes and represent a prerequisite for developing advanced tools for statistical production, enabling high quality statistics. Standardization and harmonization of statistical variables and code lists are essential for improvement and standardisation of processes and at the same time, they represent a base for a single standardized production line for all surveys. Electronic methods offer new challenges and opportunities to improve the efficiency of statistical processes and get high quality incoming data, reducing costs at the same time. As a tool for electronic data collection, SSO is using eSTAT system. eSTAT is an IT system for automatic generation of web forms, on-line data collection, communication with respondents, data validation and loading collected data into final observation register. All logical expressions that are frequently used in the process of data validation have been covered by an incorporated engine for generation of the validation rules in eSTAT. The statistical task is responsible for creation of web form instance for the specified survey and its reference period, conducting the collection process according to defined dates and running mass e-mailer for confirmation and reminders for respondents. All activities related to the questionnaire generation and data collection have strictly determined order and they represent highly standardized processes in the eSTAT system.

Keywords: statistical variables, statistical processes, standards, data collection, respondents

Title of abstract: Towards a standard production process in the Spanish Industrial Turnover Indices

Rafaela Mayo, Elena Rosa-Perez
National Statistics Institute of Spain (INE)

One of the five key areas identified by the ESS Vision 2020 is to "promote efficiency in production processes" and one of the objectives within this area is to "further identify and implement standards for statistical production". In the Industry Turnover Indices, INE Spain has already implemented several standards in its monthly statistical production tasks aiming at a streamlined standardized production process. These standards comprise the use of SDMX to transmit data to Eurostat following the corresponding international standard and JDemetra+ to conduct seasonal adjustment on a monthly basis. It also encompasses the description

and documentation of the production tasks is carried out under the GSBPM, and the adoption of recently proposed data and process architectures at our office. The whole production process is described and documented using the GSBPM at the third level according to a national adaptation of this model. For every single production activity, input, output, throughput, tools, documentation and unit(s) responsible are detailed and specified. The metadata reference report, which aims at documenting methodologies, quality and the statistical production processes in general, is based on the Euro SDMX Metadata Structure. For part of the production of indices, in particular, for the editing phase, we are adopting a data and process architectures recently proposed internally at INE Spain. The data architecture is based upon a key-value pair data model in which keys are composed using the system of statistical metadata. The process architecture is based on a modular organization of different processes which are being easily reused in different surveys. We describe from a bottom-up approach how these different production standards and normalized tools in the European Statistical System are specifically orchestrated to reach a standard production process in this monthly short--term business statistics.

Keywords: standardisation, SDMX, GSBPM, JDemetra

Title of abstract: Working time, underemployment and overemployment: two different data sources with contradictory results

Martina Rengers
Federal Statistical Office FSO, Wiesbaden, Germany

According to German FSO's microcensus results for 2015, well over 2.7 million employed people aged 15 to 74 years wanted to work more hours, while 1 million persons in employment wanted to work less. This means that the number of people willing to work more hours was significantly higher than that of people wanting to work fewer hours. The underemployed, on average, would like to work an extra 11.3 hours per week, whereas the overemployed would like to reduce their working time by an average of 11.0 hours. For the same reference period the German Institute for Economic Research DIW Berlin, using a different measurement approach reported contradictory results: According to the so called SOEP, the German Socio--Economic Panel study at DIW Berlin, just under 5.3 million employed people aged 18 to 64 years wanted to increase their working hours whereas 18.0 million wanted to reduce their hours of work. On an average there are wishes for more work of 9.0 hours and for less work of 7.8 hours. As the conflicting results repeatedly led to critical press coverage, in late 2016, the DIW and the German FSO launched a joint analysis project to investigate the underlying reasons for the obvious discrepancies. The paper will present results from the analysis, focussing on the different elements that may impact on the measurement: The questionnaires of both surveys point out that an increase or reduction in working hours would involve correspondingly higher or lower earnings. What are the reasons for the completely different results? Can relevant key factors be identified for measuring working time and working time preferences?

Keywords: working time, working time preferences, measurement problems, household surveys

Title of abstract: Electronic data collection in the Hungarian agricultural statistics

Péter Tóth
Hungarian Central Statistical Office

In year 2016 the Hungarian Central Statistical Office (HCSO) focused on two main challenges: in summer Farm Structure Survey (FSS) was carried out, while the sample-based population enumeration, called Microcensus was carried in autumn. In the beginning of 2015 it was decided that these key surveys have to be carried out without using paper questionnaires. Two kinds of applications were developed: one for the online period, and one for the interview period. In case of FSS more than 15 percent of the data provider private holdings completed the questionnaire online out of those who had opportunity for this. The enumerators used almost the same application during the interview period. Only a small part of them used HCSO-owned tablets, the others were using their own devices. In case of the FSS the application included more than two thousand controls within and between tables: one part was running during the data entry and the others before the finalisation of the questionnaire. The questionnaire could be sent after all errors corrected. The system contains a monitoring which is a web-based application that allows to maintain the pre-loaded address list, monitoring the field work and questionnaire processing process, perform payment and reporting. Instead of wasting time on data entry after the survey, more time could be spent on quality control. Using electronic devices and e-questionnaire allowed that the data processing was faster and the results could be published two months earlier than three years before. The budget were less by 20 percent compared to the FSS 2013 since there were no cost for the data entry, and 93 percent less paper was used. Beside the mentioned advantages of this innovation, the environment was also winner since cca. 160 trees were saved with using only a small amount of paper.

Keywords: agricultural survey, electronic devices, quality

SPEED TALK SESSION 10 – ESTIMATION AND CALIBRATION

Title of abstract: Effects of the influence of calibration procedures on reliability of indicators estimated by the state household sample surveys in Ukraine

Volodymyr Sarioglo
Ptoukha Institute for Demography and Social Studies, Kyiv, Ukraine

Procedures of statistical weights calibration are increasingly applying at the indicator estimation stage of the state household sample surveys in Ukraine. The practical purpose of calibrated estimation is to coordinate the results of surveys with the actual auxiliary information. This allows, on the one hand, to somewhat reduce the possible bias of estimates, and, on the other hand, to ensure matching of estimates with data from other sources. At the same time, calibration can lead to a significant decrease of the effective sample size and, accordingly, to a decrease of the accuracy of estimates. For instance in the State Household Living Conditions Survey (HLCS) provided by the State Statistics Service of Ukraine on quarterly basis, statistical weights are calibrated using relatively large number of external sources. This is primarily the data of demographic statistics on the number and structure of the population, data on the number of households, including households with children. In the experimental mode, the national accounts data and the tax administration data are used. Taking into account that the HLCS is the main source of information for measuring of a number of important indicators which in details reflect incomes, expenditures, consumption features, poverty of Ukrainian households it is important to estimate the impact of calibration on the reliability of key indicators. The article presents materials of estimation of influence of various calibration procedures on the quality of the statistical weights and on the reliability of indicators. Recommendations are given to optimize calibration procedures with the aim of minimization of their negative impact on reliability of estimation.

Keywords: calibration, sample survey, reliability

Title of abstract: A reweighting method to adjust for enterprise changes: the Italian quarterly job vacancy rate and hours worked estimates

Diego Chianella, Annalisa Lucarelli, Marina Sorrentino, Emilia Matera, Giuliano Latini
Italian National Institute of Statistics (ISTAT)

The Italian National Institute of Statistics produces quarterly estimates of the job vacancy rate and hours worked for EU Regulations and national dissemination through calibrated weights. Job vacancy and hours worked data are collected by two direct business surveys. The survey samples are drawn from the Italian statistical business register (ASIA), updated to two years before the reference quarter of the survey data, while the calibration constraints are derived from an administrative based source for the reference quarter. The distance in time between the reference periods of sampling frame and source for calibration constraints implies that sample units may be classified as belonging to different strata depending on whether the information in the sample frame or in the source used for calibration constraints is used, due to changes in the enterprises' economic activity and/or size between the two reference periods. As a consequence, if the initial weights are based on the sample frame units classification, calibrated weights are found to show a high variability within some calibration strata. In particular, in a given calibration stratum, weights can be concentrated only on few sample units while the remaining ones can be near to zero. To ensure a more ho-

mogeneous distribution of the calibration weights within strata, the initial weights have been recalculated, updating the sample units classification with the information used in the calibration strata construction and based on the source used for the calibration constraints. Furthermore, the initial weights are calculated so as to correct also for non-response by the inverse of the response rate in the calibration stratum. The first empirical results obtained applying this method have shown a reduction in the estimated sampling variance with respect to those obtained using initial weights based only on the not updated information of the sampling frame.

Keywords: calibration, reweighting, short-term, business statistics

Title of abstract: Calibration as a tool to enhance accuracy and coherence in tourism statistics

Carlos Fullea Carrera, María Izquierdo Valverde, Jorge Polo Fernández
National Statistics Institute of Spain (INE)

Calibration is a reweighting method widely used to improve the accuracy and coherence of official statistics. The Spanish Residents Travel Survey has as its main objective the measurement of the number of trips made by the resident population each month. Travelling is an activity performed by a small percentage of the population, making sampling highly inefficient. Besides, those making trips are more difficult to contact. On the other hand, the Hotel Occupancy Survey also measures the trips and overnight stays of residents staying at hotels and similar establishments, based on the information provided by the hotels. Each survey provides different figures for very similar magnitudes, creating confusion in the measurement of the tourism sector. This paper aims to describe I) the measures applied to improve the efficiency of the sample (subestratification based on a logit model predicting the probability of travelling), II) the calibration to sociodemographic variables to improve the accuracy of the estimates, and III) its extension including the hotel overnight stays provided by the Occupancy Survey in the calibration enhancing the coherence of tourism statistics. This paper assesses the differences in the main survey estimates (trips, overnight stays and average length), when comparing three scenarios: no calibration, calibration with sociodemographic variables, and calibration with all the variables, including overnight stays. In this process, regions with fewer records have to be grouped together each month due to sample size limitations. To calibrate each region separately, a reference period of three months could be considered, thereby increasing their sample size. The paper also explores the results of calibrating the three months of each quarter altogether, comparing them with the monthly estimates.

Keywords: Calibration, accuracy, coherence, tourism statistics

Title of abstract: Prediction Error Estimation of the Survey-Weighted Least Squares Model under Complex Sampling

Retha Luus – University of the Western Cape, Cape Town, South Africa
Ariane Neethling, Tertius de Wet – Stellenbosch University, South Africa

Many large-scale surveys make use of a complex sampling design where each observation unit is assigned a sampling weight which is developed over different stages. Survey-weighted least squares modelling (SWLS), the linear modelling of a continuous response based on its relationship with a number of covariates, correctly accounts for this complex sample design. One of the objectives of statistical modelling is the prediction of a future response. As such it is of importance to determine how well the selected model performs in the prediction of a future response. Cross-validation and resampling methods have long been used for this purpose under i.i.d data modelling, but not for the modelling of CS data. This talk introduces cross-validation and resampling methods for the evaluation of SWLS models based on the model's prediction error. The investigation of the performance of the different prediction error estimation methods are evaluated through a simulation study. The Income and Expenditure Survey 2010/2011 of Statistics South Africa will form the basis of the analysis. The simulation study will also investigate whether the SWLS model's predictive performance is improved through the truncation of outlier sampling weights. For this purpose two new thresholds, viz. the 1.5IQR and Hill, will be introduced.

Keywords: Cross-validation, Bootstrap, Sampling weights, Benchmarking, Trimming

Title of abstract: The R-Package 'surveysd' – Estimating Standard Errors in Complex Surveys with Rotating Panel Design

Johannes Gussenbauer, Alexander Kowarik, Matthias Till
Statistics Austria

There is an urgent need for regional indicators, especially on poverty and social exclusion, by policy makers in the EU. Surveys which were designed to estimate these indicators on national level, such as EU-SILC, usually do not provide the required precision for reliable estimates on regional levels like NUTS2 and below. With the R-Package ‚surveysd', we present a package to improve precision and estimate standard errors for social indicators on regional levels in a straightforward way. Regional estimates from subsequent waves are simply cumulated over time, assuming that these structural patterns remain fairly robust. Variance estimation for pooled data is complicated due to a high correlation within the pooled data. The package resolves the problem by using bootstrap techniques that incorporate pooling of correlated samples, like annual waves of EUSILC. In addition to variance estimation for point estimates the variance estimation for differences is supported. Usability of the package and variance improvement, using this bootstrap methodology, is demonstrated on EU-SILC UDB-data of selected countries with various sampling designs.

Keywords: Variance estimation, bootstrapping, EU-SILC, R-Programming

Title of abstract: Using administrative registers to improve sampling of EU-SILC in Austria

Thomas Glaser, Johannes Gussenbauer, Alexander Kowarik
Statistics Austria

EU-SILC in Austria uses income register data from administrative sources that are linked to the sample on micro-level to collect information about most components of the household income. Since this information is also available for the sampling frame it can be used for the sampling design. If the main research variables, especially the rate of people at risk of poverty or social exclusion (AROPE), were known beforehand, stratifying the sample by those variables would reduce the standard error of these variables and thus creating a more efficient sample. Since the former is obviously not the case, even stratifying by variables that are correlated with the main variables of interest should help reducing the standard error. For each person registered at addresses of private households in the sampling frame of EU-SILC 2016 and 2017 income data from registers were matched. The sum of all net income components were then aggregated for each address in the sampling frame creating a net household income based solely on register information. The first quartile of this characteristic is correlated with AROPE. Therefore it was decided to use it as a stratification criterion for the selection of the first wave sample of EU-SILC 2016 and 2017. Results show indeed a reduction of the standard error of AROPE and other main indicators, but only if other rotational samples of EU-SILC are not taken into account too. For the first wave of EU-SILC 2018 this approach was enhanced by using the newly available rich frame. It is a quarterly generated frame of the whole Austrian population based on a couple of different registers. Socio-demographic variables in combination with the available income information were used to train a machine learning model for estimating the AROPE for the whole frame. This predicted AROPE was then used as stratification variable.

Keywords: EU-SILC, sampling, register data, rich frame, machine learning, AROPE, predicted AROPE

Title of abstract: The Bundesbank's Research Data and Service Center (RDSC): Gateway to Treasures of Micro-Data on the German Financial System

Rafael Beier
Deutsche Bundesbank, Frankfurt, Germany

The Bundesbank is one of the largest data producers in Germany and the data are of high quality because they are quality-tested administrative micro data. Due to legal requirements and in order to meet data protection requirements, individual data can be made available only under certain restrictions. In 2014 Bundesbank set up the Research and Data Center (RDSC) which provides analysts and internal and external researchers with access to selected micro-level data in the context of independent scientific research projects.

Some of RDSCs main tasks are:

Provide data access and data protection.
Mediation between data producers and external users.
Responsibility for the methodological improvement, physical provision and comprehensive documentation of high-quality microdata sets.
Consultancy and support services to prospective and existing data users inside and outside of the Bundesbank.

Provision of standardized linked micro data sets containing multiple data sources, not only from internal Bundesbank micro data but also from external data providers.

In order to create useful and comprehensive meta data for the provided micro data, global standards like SDMX and DDI are used. For the linkage of different micro data sources, a record linkage group creates high quality master data sets that enable the linkage not only between multiple Bundesbank micro data but also micro data provided by external services like company data or else. Consultancy and support services are provided by experienced data and research experts in the areas of economics, finance and social studies.

Keywords: Micro data, research data center, data provision, financial data

Title of abstract: Improving Monthly Estimates of Job Vacancies Survey With Statistical Model

Daniel Roash
Central Bureau of Statistics Israel

The monthly Job Vacancies Survey in Israel began in 2009 by the Central Bureau Of Statistics. The survey's goals are: To serve as a leading indicator for the cyclicality in the labour market (during recession firms will begin by reducing their job openings and only later proceed with dismissing employees). To aid in assessing the demand for labour and identifying work opportunities by industry and composition of employed persons. To supply a broad view of the labour market by comparing estimates of job vacancies and the profile of workers requested by employers (labour demand) against the estimates of job seekers and their profiles as derived from labor force survey data (labour supply). In order to reduce the response burden, it was determined that the firms that belongs to the take some strata (small and medium firms) will be divided randomly during the quarter. While the large firms from the take all strata will report monthly. The strata are determined by the number of employees in the firm while the survey's target is the number of vacancies in each firm. Sometimes those variables (employees, vacancies) do not have high correlation (there is a large variation of the vacancies between firms in the same strata). This sample design led to high variation of the monthly estimates with an error pattern and a high correlation to the estimates in period t-3. We applied a State Space model that reduces the sample error and smooth's the estimates. This process improved the quality of the survey's estimates and I will present this method which is strongly relevant to surveys with panel sample. The presentation will also analyze the unique characters of the vacancies survey and it implication to imputing missing items.

Keywords: quality, model, imputation, vacancies

POSTER SESSION 1

Title of abstract: A method on reducing small sample bias of the obtained as a matrix product Gini Index

Eleni Ketzaki, Nikolaos Farmakis
Aristotle University of Thessaloniki, Greece

Gini Index is one of the most commonly used measure of income inequality. It has been noticed that it is a downward biased measure in small populations. In this study we propose an expression that aims to reduce bias, in case that the Gini Index is interpreted as a matrix product. First of all we examine the bias correction for a number of individuals and then we investigate the bias due to grouping. The formulation of the Gini Index as a matrix product has been chosen because it allows easier and faster computation. Moreover the calculation of the Gini Index based on data grouped by categories, commonly arises with income data, that are usually grouped for confidentiality purposes. In the end of our study we provide a numerical example in order to compare the proposed method with the one that already exists regarding to the bias correction.

Keywords: Gini Index, small sample bias, decomposition of the Gini Index

Title of abstract: Procedures to ensure the quality of HICS survey data in Egypt

Mostafa Younes
Central Agency for Public Mobilization and Statistics (CAPMAS), Egypt

The concept of data quality covers several aspects, from the planning of the survey to the end of it, dissemination and utilization of the data. The most important components of statistical quality elements are: accuracy, comparability, and quality control procedures. Data accuracy includes multiple aspects of any survey, could divided to sampling errors and non-sampling errors. Non-sampling errors returns to the way the work is performed and the survey tools. The national statistical office of Egypt (CAPMAS) conducting the household income, expenditure and consumption survey (HICS) survey every 2 years Started from 2008, which aim to better monitoring the changes in standard of living of the family. The present survey 2017 is Thirteenth in the series of these HICS surveys. Wherefore, there are a number of actions taken to reduce non-sampling errors in the household income, expenditure and consumption survey (HICS). The paper will review the Quality control procedures carried out on the survey, starting from the selection of the fieldworkers, field work training, income and expenditure unit duties, call back unit, the field work quality team and field work technical team. The household income, expenditure and consumption survey (HICS) 2015 response rate was 95.9% when the sample was 25000 households and was designed to be representative at the governorate level, have been visited for two research periods.

Keywords: Non-sampling error; Response rate

Title of abstract: A novel approach to detecting pattern changes in time series from Eurostat

Álvaro Gómez Losada, Panayotis Christidis
Joint Research Centre - European Commission, Spain

Detecting pattern changes in time series (TS) is of the utmost importance when evaluating a phenomenon's trend shifts over time. In particular, studying these changes on the latest observations of any TS can help to inform decisions as to what actions to take on a particular issue. The methodology for formalizing detection of these changes at the time an official estimate is released for a TS should be based on well-settled, appropriate methodology. This study introduces a new method based on Hidden Markov models (HMM) to ascertain whether the latest figures of a TS might be experiencing a shift or structural break with respect to the regular trend. Ultimately, these latest TS observations are the most important for policy makers. HMM were used to detect regimes (clusters) in TS and to characterize them using their average and standard deviation values. These latest observations of TS are considered not to experience a shift when their values are within a confidence interval (CI) estimated using those average and standard deviation values from their regimes. This CI can be modified to detect when the last observations are experiencing a shift or structural break with regard to the TS trend. This methodology is modular and flexible, allowing any TS to be studied according to their characteristics over time. The R software was used to deploy this methodology, and TS with different lengths and characteristics from Eurostat were tested. Results obtained show promising sensitivity results in detecting shifts or structural breaks in TS.

Keywords: Hidden Markov models, time series, Eurostat, nowcasting, open data

Title of abstract: Improving the quality and relevance of the official statistics produced in the SSO of the Republic of Macedonia

Natasha Cenova
State Statistical Office, Skopje, Macedonia

This paper, in the context of improving the quality of collected data on issued building permits, analyses the quality and relevance of the official statistics. Data on issued building permits are collected according to the Eurostat Methodology of short-term business statistics, Council Regulation (EC) 1165/98, Commission Regulation (EC) 1503/2006. The indicators for issued building permits are part of the PEEI main economic indicators for the economic developments in the EU. They also provide important information on the economy and monetary policy as well as for future investments in the construction sector. In the Republic of Macedonia, the survey was implemented in 2008 with paper questionnaires sent to the authorities responsible for building permits. In 2014, the Government of the Republic of Macedonia started a project for online data collection via an electronic system for issuing building permits, e-permits, prepared by the responsible authorities - the Association of the Units of Local Self-Government (ZELS) together with the Ministry of Transport and Communication. In cooperation with these institutions, the State Statistical Office incorporated the variables needed for the survey by adding a questionnaire for the official statistics. The reporting units provide information on the required variables directly on the e-permits website while applying for a permit. In this manner, the quality and coverage of the collected data from administrative registers is improved and the relevance of the official statistics is raised to a higher level.

The published data on issued building permits are used by the Government, the business community, international organisations, etc.

Keywords: relevance, official statistics, building permits, PEEI, ZELS

Title of abstract: The ESA 2010 Pensions Table: an integrated view on the functioning of pension systems in Spain

Sixto Muriel de la Riva, Carlos Javier Valero Rodríguez, Andrés García Carreira
National Statistics Institute of Spain (INE)

The inexorable impact of the population ageing, the peculiarities of pay-as-you-go schemes of public systems and the increasing role played by private systems in developed societies emphasize the need of a harmonized measure of right and obligations accrue in them as one the main priorities for the statistical systems. Current national accounts standards (SNA 2008 and ESA 2010) already include guidelines for the registration in their systems of all private pension obligations/rights linked to employment regardless of whether they are systems with or without constitution of reserves. In addition, they propose the recording of all pension schemes, including contingent obligations/rights accrued in public systems in an annexed table. The supplementary table for pension rights accrued to date in social insurance schemes will allow us to see the evolution of all pension rights stocks and the flows that motivate their variations, regardless the fact they are non contingent financial assets/liabilities for the households/pension managers or not. Both the objectives and data compiled in the table present obvious conceptual difficulties and require a high level of expert knowledge in the financial, insurance and actuarial fields. Thus, in the Spanish case, the close collaboration with external agencies from various areas has been a basic component of the project, as a clear example of interinstitutional cooperation towards the highest standards of quality in official statistics. Furthermore, a didactic dissemination of the pension tables results as a tool for analyzing the functioning of national pensions systems but not as a measure of their future feasibility is one of the most challenging issues that ESS and other international organizations face nowadays.

Keywords: pensions, national accounts, ageing

POSTER SESSION 2

Title of abstract: Compilation of the House Price Index in Cyprus

Costas Diamantides, Athanasia Petsa
Statistical Service of Cyprus

The Statistical Service of Cyprus has applied the Rolling Window Hedonic Regression model for the compilation of the Cypriot House Price Index. This method was selected in order to quality-adjust for the characteristics of the dwellings. The methodological work for the revision of the compilation method has been combined with improvements in the methods of data collection in order to increase the accuracy of the index. This paper describes the peculiarities of the Cyprus residential market leading to the specific form of the model. An important conclusion is that a „luxury-house" market with foreign buyers constitutes a significant part of the Cyprus Residential Market. Additionally, the improvements of the data collection methods, leading to data of better quality, are presented. The main contribution of the new collection methods is that additional characteristics of the dwellings can be used. The method of compilation of the House Price Index is explained. The stratification scheme of new and existing dwellings, the explanatory variables introduced in the model in order to quality-adjust for the characteristics of the dwellings and the weighting scheme of sub-indices are described. The regression coe fficients, the coe fficient of determination R2 and the plots of residuals, presented for contracts of sales for a specific quarter, show a good fit for the model. Finally, the paper discusses areas for future improvements on the method of compilation of House Price Index.

Keywords: House Price Index; Laspayres Index; Rolling Window Hedonic Regression; Transactions of Residential Dwellings.

Title of abstract: The importance of cognitive interviewing in quality official statistics: The case of smoking habits

Gaja Zager Kocjan, Darja Lavtar, Metka Zaletel
National Institute of Public Health, Ljubljana, Slovenia

Providers of official statistics increasingly recognize a critical role that qualitative pre-field testing has in identifying problematic survey questions. Pre-field methods include cognitive interviews that focus on investigating mental processes involved in answering questions. At the Slovenian National Institute of Public Health, this in-depth qualitative method was employed to thoroughly examine a question on smoking habits, followed by a quantitative testing of the original and revised question in the pilot study. The analysis of cognitive interviewing data revealed the following problems: (I) it was unclear whether all smoking-related products or just tobacco products should be considered, (II) several occasional smokers did not regard themselves as current smokers but rather as former smokers or non-smokers, (III) respondents that have smoked only for a short period in their past regarded themselves as non-smokers or former smokers. Based on these findings, the question on smoking habits was revised: a note was added explaining which products should be considered and two additional answer categories were included, covering occasional smokers and non-smokers that only tried smoking few times. In next step, a pilot study was conducted using computer assisted web interviewing (CAWI) with a simple random sample (SRS) of 600 inhabitants splitted into two halves (the final samples consisted of 82 and 75 respondents), allowing to test and compare the original and

revised version of the question. The findings revealed a greater proportion of current smokers (regular and occasional) and a lower proportion of former smokers in a sample answering to the revised question. The proportion of non-smokers (those who never smoked or only tried smoking few times) was largely similar in both samples. Altogether, these findings confirm the importance of cognitive interviewing in identifying problems regarding understanding and answering survey questions, particularly those concerning sensitive topics, prior to collecting data in the field.

Keywords: qualitative pre-testing, cognitive interviews, pilot testing, health-related measures, smoking habits

Title of abstract: Are we on the same page? Ensuring the synchronisation of data across multiple platforms

Colm Bates, Chelsea Murray
European Central Bank, Frankfurt, Germany

It is important for agencies using official data to ensure the data displayed on their websites and in current publications match those published in official source databases. Ensuring such aligns with the principle of accessibility and clarity and is necessary for institutions holding a commitment to data quality. Data transmission infrastructure, including transmission standards and data taxonomies facilitate the transmission of data and subsequently synchronisation of data published by different organisations. However, unobserved obstacles in the data dissemination, transmission, and production processes may occur and hinder the synchronisation. Institutions relying on out of date or incorrect data face damage to their credibility. Even if policy makers use the correct official figures in making decisions, it is important from a public perception point of view that the data published on their website match official figures. We demonstrate two tools to facilitate checking the synchronisation of data. The first compares the data published on the Statistical Data Warehouse of the European Central Bank (ECB) with those officially published by Eurostat. The second tool automatically reads portable document format (PDF) tables prepared for ECB publications and compares the data presented with the source database.

Keywords: Data synchronisation, Data dissemination, Process automation, Data quality checking

Title of abstract: Impact of applying quality standards on statistical products

Haidy Mahmoud
Central Agency for Public Mobilization and Statistics (CAPMAS), Cairo, Egypt

This paper aims to improving the quality of statistical products in Egypt through to examine the impact of applying the total quality standards on household survey, focusing on „the income, expenditure and consumption survey" as one of the most important field survey carried out by statistical agencies worldwide, true to the life of the Egyptian citizen. The study used descriptive analytical and quantitative statistical methods, based on the evaluation of the statistical status of Egypt, the European code of best practices, evaluation of experts and citizens on the quality of statistical data, the use of the statistical capacity index as a good means of comparing the performance of the Egyptian statistical apparatus with its counterparts in the Arab, African and international countries. In Egypt, in addition to the SWOT analysis. Statistical methods were used to identify the correlation between the various variables. The study showed that the standard

(accuracy) is the most affect by 38% correlated with the other quality standards; then the (availability) is 21%; that both the standard (accuracy) and (availability) affect 59% on the quality of the output of the statistical product, It was also found that 99% of the sample frame design affects the quality of the statistical product, 61% of the response rates in the field work are due to the accuracy and clarity of the statistical form used in the data collection form, only 7% of the researcher's training on the data collection form affects the fieldwork method, it was found that 55% of the analysis of the survey results is due to the accuracy of published data.

Keywords: Household Survey, Statistical Capacity

Title of abstract: The Istat platform prototype for the documentation of administrative data used as input of the official statistics

Grazia Di Bella, Marina Venturi
Italian National Institute of Statistics (ISTAT)

Assessing the usability of administrative data (AD), measuring their relevance for the statistics production process, monitoring the AD collection, document the AD treatment are the main objectives of the QRCA (Quality Report Card for AD) internal web site that Istat is going to implement. Many Process and Structural metadata are available directly from the IT tools supporting the AD collection and treatment processes already in use: a) the ARCAM system that provides an interface for the AD owners and a tool to manage collection (150 Administrative sources) and guarantees the security of the data transmission process; b) SIM (System of Integrated Microdata), the DB (Database) of integrated administrative microdata (70 Administrative sources); c) the National Statistical Programme DB reporting statistical uses and making AD accessible for statistical purposes in in terms of compliance with the data privacy law. Once the interoperability of the systems is ensured, these metadata, together with some additional reference metadata, will make availableto internal users to support their statistical production processes. The quality framework used to report (derived from the international project BLUE-ETS), is composed of three Hyperdimensions: Source, Metadata and Data. In each Hyperdimension, dimensions and quality indicators are considered. In the Source Hyperdimension the quality aspects related to the source as a whole are reported. The Metadata Hyperdimension considers conceptual aspects such as the administrative objects (i.e. events and units) and variables and the Data Hyperdimension related to the data quality (facts) includes quality indicators applied to data. For the realization of the prototype the tool that is being used is the Business Intelligence Visual Analytics Platform - MicroStrategy adopted by Istat to realize some data visualization products. Supported by a Java application and a passing Oracle Database, it allows to produce tables and graphs and to generate various types of output.

Keywords: administrative data, quality, metadata

Title of abstract: Bringing statistics to users and users to statistics

Jose Jabier Zurikarai, Cristina Prado
EUSTAT, Spain

Statistics are not always as useful, accessible and easy as we would like them to be. The gap between Official Statistics and our users is wide and should be tightened. There are high lever users which can have a fluent relation with us, but what about the rest of them? Students, journalists, local authorities, municipalities and govern department employees and others are not always very familiar with official statistics; sometimes due to their short experience and other times because they are random users. World has gained an increasing complexity for everyone, so we should be aware of the real situation of users. In this context, the achievement of quality framework principles, such as accessibility and clarity, needs some real action. We have to meet the users, talk to them, ask for their expectations, explain the value of our data, how to use data, teach them to be efficient and so on. We believe real interaction has to take place and there are two options: they come to us, or we to go to them. Indeed we are doing and planning both ways. The aim of the paper is to share our experience and future projects. We regularly go to the universities, to the classroom, and teach them what official statistics are and how to use them. We also do similar activities with journalists and public employees. But this is time consuming, so we plan to develop e-learning programs to have large audiences. Activities involving direct interaction are also very valuable because we get some feedback from those students who will be our future high-end users. We will explain what we do and how we do, but also the problems found and the solutions in the making, suitable activities for any agency.

Keywords: statistics in the clasroom; collaboration with teachers; workshop with professional users

Notes


100 lat GUS
Statistics Poland


eurostat