Annex I - Description of the action

1.1. Description of the Work

The ESSnet on SDMX phase II is a collaborative network between partners of the ESS and aims to accelerate the implementation of the standard SDMX at national and international level (see technical specification of the Action). The harmonization on the provisioning and dissemination formats of data and metadata sets improves accessibility and clarity in the access of statistics, specially for the consumers of statistics.

To reach this goal it is essential to elucidate and motivate countries on the benefits they can get by sticking to architecture that enables the exchange of information on a standardized and smooth way.

SDMX is an international standard sponsored by seven international Organizations among them Eurostat, and, in this context, can been seen as an efficient “tool” for harmonizing, exchanging and sharing statistics within the ESS. But to achieve this issue SDMX must be introduced not only like a new data and metadata format, but like the backbone of the business processes behind each statistical information system.

In fact SDMX besides defining data and metadata file formats, is based on an information model and technical architecture for a more efficient way to exchange and sharing statistical information.

The architecture detailed in SDMX is based on SOA (Service-oriented architecture), where web services approaches are used to minimize the cost of migrating between different systems based on “loosely coupling” applications using XML interfaces. Typically these interfaces are based on SOAP or REST web services and applications only need to agree on the XML format to interoperate. Cost of new development is lower than with traditional “hard-coded” APIs (“application program interfaces”). Any application which understands the XML can be “swapped in” to replace an existing component.

However “SOA” works well between disparate systems implemented in different organisations but once inside a specific system, the system must work with program code. Modern systems use object oriented (OO) programming languages such as Java and C#. Many organisations are writing SDMX systems using OO techniques and making their source code available to others. This is a step forward but it is not enough: different systems have different approaches to the granularity of the functions that are supported. The SDMX Information Model helps but in itself it is not sufficient to ensure interoperability of different OO components So it is not possible to “mix-and match” parts of one system with parts of another system. Moreover OO components developed by different subjects can follows different methodologies in designing, implementation and documentation.

These factors don’t facilitate the re-using of the software. This negative impact is at two levels:

a. statistical Organizations or private Subjects, that start to develop SDMX software, hardly ever decide to develop new needed functionalities upon an existing software. This due to the difficulties to handle what has been developed by others: very often the already existing software is not well documented and/or to enter into its “logic” is not an easy task. The consequence is the proliferation of software that implement the same set of basic functionalities (this basic functionality could cover also the 80% of the needs), more few different functionalities;

b. when a statistical Organization decide to invest in SDMX has difficulty to choose, among the several alternatives, the software that will fit the needs. A real use-case is when a statistical Organization is involved in two or more SDMX projects. In this case very often it is necessary to re-use two different software developed with different technologies, documented in different ways and hard to connect each other. Paradoxically the re-use of the software could introduce new complexity within an existing Information System, and consequently increase the so called stove pipes.

In the last two years, in order to stimulate the use of SDMX within the ESS, Eurostat has started the development of an SDMX infrastructure that countries can re-use to speed the implementation of SDMX as common standard for sharing data and metadata, and harmonizing statistics.
At the beginning the main aim of the development of the "SDMX NSI Reference Infrastructure" was to provide the countries with a non-invasive infrastructure that could "wrap" existing statistical dissemination/reporting architecture and make them SDMX compliant. Now that the "SDMX NSI Reference Infrastructure" is available and enough mature, the next step is to foster a community of developers from the National Statistical Institutes and other statistical Organizations with the main aim to encourage further developments based on that infrastructure, so to provide to the SDMX community a set of software that can be re-used partially or as a whole, but that they can guarantee homogeneity and interoperability.

The ESSnet on SDMX phase I, which will finish at the end of March 2011, has already adapted some of its objective, in order to fit the aims, described above, of the "SDMX NSI Reference Infrastructure" produced by Eurostat. Two representative examples are documented in WP4 (Generalized SDMX Framework) and in WP7 (PC-Axis with SDMX-ML).

The ESSnet on SDMX phase II moves a step forward and besides continuing some of the work of a previous ESSnet on SDMX phase I (see WP2 and WP3), wants facilitate the rationalization of the current offer of the SDMX software in order to foster both a community of SDMX developers and users. In this context the SDMX NSI Reference Infrastructure, more than the ESSnet phase I, will have a central position in all the tasks.

This action will last 18 months, and the partner countries are: Italy, Slovak Republic, Sweden and Bulgaria. It'll coordinated by Italy. All the Organization are statistical authorities. In the case of Slovak Republic, the statistical Office made a statement declaring that "INFOSTAT is a National Authority responsible for research in statistics and participates in development, production and dissemination of official community statistics. Therefore it is eligible to be a beneficiary of the ESSnet project". This statement as well as the "Deed of Foundation" of INFOSTAT can be met in point 2.12 of this proposal.

The Action is composed by 5 work packages (WP). The Management WP has only one responsible, the co-ordinator country. Below a description of the WPs.

**WP1 – Italy**

Assure the overall management of the Action, thus covering all the aspects related to an efficient management of the action and consequently of each WP, resources, schedules, issues and adequate implementation. The budget of this WP encloses four travels to Luxembourg for all partner countries, with the porpoise of the co-ordinator meeting.

**WP2 – Sweden**

The aim of the work package is to integrate PC-Axis with SDMX continuing the work started in the first ESSnet on SDMX. The results could be provided for the PC-Axis community.

**WP3 – Slovak Republic**

The aim of the work package is to support the users that are responsible for survey processing at NSI in their effort to implement SDMX standard in the key phases of the statistical business process. The work package activities will resume the results achieved in the ESSnet on SDMX project Phase I, mainly on the results achieved in the WP8 oriented on the implementation of SDMX in survey preparation and data collection processes. The activities will comprise of methodological tasks and development of IT applications for DSD generation/transformation. The DSD generating will support the SDMX implementation in the phase that follows data collection and will also support the eventual dissemination of statistical micro data in standard format. A pilot implementation of the application on selected statistical domain will demonstrate a usability of the SDMX standard in these particular phases of the statistical business process.

**WP4 – Italy, Bulgaria** (Italy and Bulgaria are both taking part in this WP, Italy is responsible)

Several tools, software and entire framework are already available within the SDMX community as open source packages. From the very first beginning the SDMX sponsor Organizations have been fostering the development of software open source, in order to facilitate the re-usability. Software re-usability is one of the main benefits that a statistical Organization can consider to decide in investing in SDMX. Unfortunately it is very difficult to choose, among the several alternatives, the software that will fit the needs. At this purpose on the SDMX web site, an instrument for cataloguing and categorizing the available software was insert. But this is not sufficient, in fact in many meetings a recurrent topics are based on how to share and re-use software. Very often when a statistical Organization...
Role: The SG will

- Define the project;
- Analyze monthly reports;
- Authorize changes in scope, timescale or budget of the Action if significant changes are predictable;
- Evaluate quality and approve deliverables.

Working method:
- Continuous contacts (can be e-mail, audio or video conference), among Eurostat, co-ordinator and partner countries.

Coordinator:

Italy

Role: the main responsibilities of the coordinator will be on planning, follow-up and control of the action, namely it will:

- Prepare the Action plan with the contributes of the co-partner countries;
- Supervise compliance with the approved plan;
- Maintain continuous contact with Eurostat and with partners countries;
- Produce monthly progress reports and forecasts of the Action to the Steering Groups;
- Make monthly progress reports available on the action side, after SG approval;
- Report to the SG significant deviations forecasted or already occurred;
- Overseeing the administrative and financial coordination
- Timetabling management and project team meetings.
- Ensuring fulfilment of contractual agreements, both by the overall project and by individual partners.
- Coordinating overall work-packages.
- Monitoring project progress with respect to quality and completion of all deliverables, workplan milestones and vision.
- Ensuring high quality of the deliverables.
- Setting-up and maintaining a wiki for structured document repository and project communication (if Eurostat provides this facility), linked to the project website, with unrestricted access for project partners.
- Overseeing dispute resolution.

Working method:

- The coordinator country will maintain continuous contact with Eurostat and with other partner countries;
- The monthly progress reports, record of changes and record of risks will be produced based on the WP’s monthly reports.

Partner country:

Role: Each work-package has a WP leader, who is responsible for the production of the deliverables of each individual work-package. Not all the countries are leaders of WP

- Supervise compliance with the approved plan;
- When the country is responsible of a WP, coordinate the technical work of WP;
- Prepare eventual changes to the WP plan and submit them to the co-ordinator
- Report problems and risks within the WP to the co-ordinator or to the WP responsible;
- Produce monthly report and forecast to the co-ordinator or to the WP responsible;
- When the country is responsible of WP, organize and chair the WP’s meetings;
- Make deliverables available on time to the co-ordinator or to the WP responsible;
- Make documentation and deliverables available on the Action site, after SG approval;
- Provide cost statement to the coordinator or to the WP responsible;
is involved in two or more SDMX projects, it is necessary to re-use software developed with different technologies, documented in different ways and list but not last hard to connect each other: paradoxically the re-use of the software could introduce new complexity within an existing Information System, and consequently increase the so called stove pipes.

In the last two years, in order to stimulate the use of SDMX within the ESS, Eurostat has started the development of an SDMX infrastructure that countries can re-use to speed the implementation of SDMX as common standard for sharing data and metadata, and harmonizing statistics. Now that the “SDMX NSI Reference Infrastructure” is available and enough mature, the next step is to foster a community of developers from the National Statistical Institutes and other statistical Organizations with the main aim to encourage further developments based on that Infrastructure, so to provide to the SDMX community a set of software that can be re-used partially or as a whole, but that they can guarantee homogeneity and interoperability.

The WP has the objective to create, the necessary instruments that will allow the community of developers to easily contribute to the further development of the “SDMX NSI Reference Infrastructure”.

**WP5 – Italy**

Several capacity building actions are already put in place by Eurostat and other Organizations. In particular Eurostat within the ESS has been providing SDMX training sessions for Statisticians and IT Staff. Moreover a set of self learning tutorials (composed by videos, student books and self-tests) are available for downloading from the SDMX website. Finally the SDMX User Guide was updated in 2009.

The idea of this WP is to not overlap the above actions, but transfer knowledge and experiences from NSIs to NSIs.

**1.2. Organisation**

1.2.1 Governance:

The organization is composed by a co-ordinator Istat – National Statistical Institute of Italy and the following co-beneficiaries:

- Institute of Informatics and Statistics - INFOSTAT
- Statistic Sweden – SCB
- National Statistical Institute of Bulgaria – BNSI

The project co-ordinator will be responsible for the ESSnet administrative, financial, and technical management, it will ensure close cooperation between team members involved in the different tasks. The coordinator will be particularly concerned on identifying in early stage problems that might occur, to solve them and minimize the impact on overall action.

In the next sections will be detail the different responsibilities.

**The sponsor:**

European Commission

**Role: the sponsor will:**

- Authorize the plan and the budget of the action;
- Authorize the start of the action;
- Authorize the closure of the action.

**Steering Group**

**Composition:**

Eurostat and the partner countries;
Working method:

- Meeting accordingly to the plan;
- Other contacts with the WP co-partners and co-ordinator
- Monthly reports to the coordinator

1.2.2 Risk management:

There could be two kinds of risk factors: those which are inherent to the project itself and those who can occur from specifics of a particular WP and/or from particular situations of a partner country.

Therefore partner countries should evaluate very carefully, prior to the commitment to a specific WP, all the aspects (resources, technical issues, etc.) that are critical success factors, consequently the ones where the likelihood of risk is high.

A list of problems that potentially can arise is:

<table>
<thead>
<tr>
<th>WP</th>
<th>Risk</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lack of full cooperation and coordination between partners of the same WP</td>
<td>Countries will be noted that are compromising the project. Behaviour must be corrected as soon as possible. If persist country will be off the Action, no funding will be given.</td>
</tr>
<tr>
<td></td>
<td>Under estimation of resources</td>
<td>Immediately re-estimation of resources on a realistic basis.</td>
</tr>
<tr>
<td></td>
<td>Tasks improperly defined and/or scheduled</td>
<td>Redefinition and rescheduling, eventually allocating more resources to remain of schedule</td>
</tr>
<tr>
<td></td>
<td>Lately report of unexpected constraint or any other kind of issues that compromise each WP</td>
<td>Countries will be strongly advised not to do the same mistakes. Eventual rescheduling and/or additional resources if possible</td>
</tr>
<tr>
<td></td>
<td>Technical problems due to the adoption of non robust tools or insufficient knowledge necessary to an suitable development on time</td>
<td>Replacement of tools by robust as well proven ones assuming that internal knowledge exists. Eventual need of external consultancy and support</td>
</tr>
<tr>
<td></td>
<td>Lack of resources due to a internal reallocation.</td>
<td>Redefinition and rescheduling</td>
</tr>
</tbody>
</table>

1.2.3 Quality plan:

We propose a set of criteria to evaluate the deliverables of the Action:

<table>
<thead>
<tr>
<th>WP</th>
<th>Title</th>
<th>Purpose</th>
<th>Quality criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Monthly report</td>
<td>Monitor of the Action</td>
<td>timeliness, completeness, comprehensibility</td>
</tr>
<tr>
<td>2,3,4,5</td>
<td>Documentation</td>
<td></td>
<td>- Structure of the document</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Objectivity and clarity of the document</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Technical rigour of the choice done</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Usefulness of the result obtained</td>
</tr>
<tr>
<td>2,3</td>
<td>Software</td>
<td></td>
<td>- Be reusable within the European statistical system</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- robustness</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- user friendliness</td>
</tr>
<tr>
<td>5</td>
<td>Workshop</td>
<td>Capacity building</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Agreement number 60501.2010.001-2010.185</td>
<td></td>
<td>- Agenda</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Slides and others materials produced</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Evaluation survey by pfflaįafsaffiž</td>
</tr>
</tbody>
</table>
1.3. Timing of tasks

EssNet on SDMX II - Timing of the work

<table>
<thead>
<tr>
<th>Task/Work package number</th>
<th>Start Month or starting event:</th>
<th>End Month</th>
</tr>
</thead>
<tbody>
<tr>
<td>WP1 IT</td>
<td>M1</td>
<td>M18</td>
</tr>
<tr>
<td>WP2 SE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WP3 SK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WP4 IT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WP5 IT</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Objectives:
Integrate PC-Axis with SDMX-ML (continuation of first phase of EssNet on SDMX)

Description of work:
- Implementing of attributes on cell level in PC-Axis family Internet user interface PX-Web. With the attributes on cell level implemented in PX-Web it will be very easy for the NSIs in the PC-Axis community to check attributes added on cell level that has been loaded in the SQL-database or in PC-Axis files. 270 hours
- Implementation of the Generic Default SDMX-component in PX-Web. The “Generic Default SDMX-component implemented in PX-Web will give the possibility to reuse the tools for visualization using SDMX as input format to be linked into PX-Web installations. For instance the Explorer used by OECD could be used for any statistical material in the databases used by the PC-Axis community that is not using any Data Structure Definition. 55 hours.
- More use cases shall be added to the User's Manual for the Process. Addition of more types of statistics that is more complicated. Concerning more use cases for the User's Manual, it is a matter of the different complexity in different subject matter areas of statistics. It is not obvious that the solution developed in Phase 1 covers all the different statistics demanded by the Eurostat. 480 hours.

- Presentation and information for other countries and Swedish authorities that will use the process developed in Phase 1. Take actions solving problems that arise when different countries start to use the process. Countries in the PC-Axis community that would like to use their Dissemination databases to meet the demands for SDMX pull or push method from international organizations. The interest has been very large on this on the annual meetings in the PC-Axis community, but special interest has been showed by Denmark, Norway, Finland and outside the community for instance Lichtenstein has expressed their interest for the solutions in connection with PC-Axis and SDMX. 350 hours.

- Installation and test of new versions of the Eurostat SDMX Reference Architecture at Statistics Sweden. 160 hours.

- Study of how other governmental bodies in Sweden will do the reporting to Eurostat in the SDMX era. The purpose with the study of the governmental organizations outside Statistics Sweden that are responsible for official statistics to find out if they intend to use the result of the project PC-Axis SDMX integration or any of the Eurostat SDMX Reference Architecture in the reporting of statistics to Eurostat. This study will give further knowledge how to support them in the SDMX era. 150 hours.

- SDMX 2.1. The routines in the process are now being based on the SDMX 2.0. The demand in Census 2011 is SDMX 2.1. Concerning the Census hub and SDMX 2.1 we have considered few hours to put up activities in the application to handle the case of use of version 2.1 in the Census. 50 hours.

- Planning, coordination meetings, report writing and administration. Number of coordination meeting are about the same as in the Phase 1 as every half year of the project, kick off and ending meeting, report writing is every month including administrative work around that. The planning process is much influenced of what happens in the project and not least what is happening around the project at Eurostat development work etc. It is an estimation from the experiences from Phase 1. 780 hours

The Phase 2 project special attention is paid to the version 2.1 of SDMX format.

**Sub-Tasks:**

**Deliverables**

**D - 2.1 Documentation.** Documents to be delivered is use cases with different type of materials. The experiences at Statistics Sweden from the present routines for delivery of statistics to International organizations is that with different type of material and different key families there is a lot of activities to find how to prepare the materials. We foresee the same type work with the adaption of different material with the mapping to the DSD:s in the SDMX platform. The documentation is are supposed to be used in the workshops planned for other NSI:s also as a part of the Phase 2 of the project.

**D - 2.2 Workshops.** We plan two workshops to be held for other NSI:s in the PC-Axis community where also other NSI can participate as observers. The purpose with the workshops is to transfer the methods and software for use of the same platform based on PC-axis and the Eurostat SDMX Reference Architecture achievements in Phase 1 and Phase 2.

**D - 2.3 SDMX functionality integrated in the web-software**

**Background**

PC-Axis/PX-Web is software to make output databases available for the public on the Internet and for distribution on CD-ROM. PC-Axis/PX-Web is used by the following countries within or related to the ESS:
**Member states:** Estonia, Denmark, Ireland, Finland Latvia, Lithuania, Luxembourg, Slovakia, Slovenia, Spain and Sweden. **Candidate states:** Croatia and The Former Yugoslav Republic of Macedonia. **EFTA Countries:** Iceland and Norway. In total 15 countries in the region. In addition the software is used in Africa, Asia and the Pacific and Latin America.

Discussions are going on to make PC-Axis Community Source software.

<table>
<thead>
<tr>
<th>Task/Work package number</th>
<th>Start Month or starting event</th>
<th>End Month</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>M1</td>
<td>M18</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Title</th>
<th>Support of SDMX application for micro data handling within statistical business process at the NSI level</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Participant id</th>
<th>INFOSTAT Slovakia</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

**Objectives**

The aim of the work package is to support the users that are responsible for survey processing at NSI in their effort to implement SDMX standard in the key phases of the statistical business process. The work package activities will resume the results achieved in the ESSnet on SDMX project Phase 1 mainly on the results achieved in the WP8 oriented on the implementation of SDMX in survey preparation and data collection processes.

The main objective of the work is to implement SDMX standard at the NSI level in the phase of statistical micro-data processing. The activities will comprise of methodological tasks and development of IT applications for DSD generation/transformation. The DSD generating will support the SDMX implementation in the phase that follows data collection and will also support the eventual dissemination of statistical micro data in standard format. A pilot implementation of the application on selected statistical domain will demonstrate a usability of the SDMX standard in these particular phases of the statistical business process.

**Description of work**

The work shall comprise the analysis of the relevant phases of the statistical business process, the need of metadata in these phases and how the SDMX standards could be used at NSI for supporting statistical micro data description and its exchange. The system will be developed as open one with the possibility to support an use of SDMX standard in the whole business process flow. The work shall result into outputs of methodological and technological nature with the pilot application to verify the SDMX use in the of micro-data description, its further processing and eventual dissemination. The version of SDMX used is conditioned by the fact, that the tools we will be using (Data Structure Wizard, Metadata editor, Mapping assistant) will be also v2.1 SDMX compliant.

**Sub-Tasks:**

T 3.1 – Management (Begin: M1 - End: M18)

- Preparation of activity plan.
- Coordination of experts work in the analytical and designing tasks.
- Monitoring of progress achieved
- Preparing progress reports
- Preparing final report
D – 3.2 User and System documentation (Available :M16)
D – 3.3 Presentation of results achieved (Available :M18)
D – 3.4 Analysis of DDI vs SDMX (Available :Month 18)

Phase I

SDMX Metadata editor

questionnaires
definition user

questionnaire
application

respondent
user

questionnaire
metadatabase

Phase II

SDMX Data Generator

Microdata DSD generator

DSO file

Merging assistance store

Data Retriever

SDMX Dataset Dissemination File

New module developed in Phase II

Standard modules of SDMX Dissemination architecture

WP3 - Support of SDMX application for micro data handling within statistical business process at the NSI level - Timing of the work

<table>
<thead>
<tr>
<th>Main Tasks / Time schedule</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
<th>M5</th>
<th>M6</th>
<th>M7</th>
<th>M8</th>
<th>M9</th>
<th>M10</th>
<th>M11</th>
<th>M12</th>
<th>M13</th>
<th>M14</th>
<th>M15</th>
<th>M16</th>
<th>M17</th>
<th>M18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Management</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis and design</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSD generator development</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preparing user and system documentation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pilot implementation of DSD Generator</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Agreement number 60501.2010.001-2010.685
T 3.2 - Analysis and design (Begin: M1 - End: M8)
- Analysis of metadata and procedures needed for DSD generation
- Analysis of user requirements for statistical micro data
- Design of the system architecture
- Design of system functionality
- Analysis of SDMX tools for suitability of their use in the implementation phase
- Analysis of standard SDMX libraries used in NSI SDMX Reference Dissemination Architecture

T 3.3 - DSD generator development (Begin: M3 - End: M14)

Description of fig 1). In Phase-I, we used SDMX metadata messages to get metadata about questionnaire. User who defined the questionnaire should fill the metadata report in Metadata Editor. That means we collected information like: “The item in third row and second column of questionnaire is integer value of length 2 characters and it is the item with reference code AQ1824.” In second Phase, we would like to use collected information about questionnaire to create a tool, which would generate DSD in order to be able to disseminate the data. That means that developed tool will take metadata and create DSD. It will automatically recognize, that specific data forms dimensions another specific data are attributes and another are cross domain measures.

- Proposal for DSD generic structure. In this step we would like to propose a way how we would transform micrometadata about questionnaire into DSD structure. What micrometadata should be in DSD dimension, which attributes, consider using of groups and cross sectional measures.
- Metadata used for micro data description
- Development of tools for DSD generation based on statistical micro data

T 3.4 - Preparing user and system documentation (Begin: M6 – End: M16)

T 3.5 – Pilot implementation of DSD Generator (Begin: M12 – End: M17)
- Preparation of two pilots with micro data from two selected surveys
- Performing the pilots
- Evaluation of pilots results
- Report on pilot results

T 3.6 – Analysis of DDI vs SDMX Standards, proposal for possible SDMX extension towards DDI. Analyses of DDI standard, analyses of possible SDMX extension. In this step we would require cooperation with SDMX architects. Both DDI and WP8 of the ESS net phase 1 are dealing with the micro data. The one of the deliverable of WP8 is oriented on a comparison of features of DDI as described in the DDI documentation with the features of the tools developed for questionnaire design and micro data collection using the SDMX standard. More specific the comparison will be oriented on the methods and functionality of both systems for descriptions of statistical micro-data (metadata), for micro-data and its metadata storing. In the ESSnet PhaseII a more detailed study of DDI functionality and of functionality of tools developed within the WP3 will be elaborated.

T 3.7 – Demonstration of the results achieved (Begin: M17 – End: M18)
- Presentation of results on the project workshop

Deliverables

There will be three essential results presented:

- a methodology for creation of DSD generic structure (this will be a summary of experiences with the metadata used in the questionnaire description and its transformation into a DSD for micro-data),
- a report on the results of the pilot implementation of the DSD generator developed in the WP3,
- a comparative study of the DDI standard with the SDMX standard in their application in the micro-data processing an/or dissemination

D 3.1 Design of system architecture and system functionality (Available :M8)
Objectives:
Several tools, software and entire framework are already available within the SDMX community as open source packages. From the very first beginning the SDMX sponsor Organizations have been fostering the development of software open source, in order to facilitate the re-usability. Software re-usability is one of the main benefits that a statistical Organization can consider to decide in investing in SDMX. Unfortunately it is very difficult to choose, among the several alternatives, the software that will fit the needs. At this purpose on the SDMX web site, an instrument for cataloguing and categorizing the available software was insert. But this is not sufficient, in fact in many meetings a recurrent topics are based on how to share and re-use software. Very often when a statistical Organization is involved in two or more SDMX projects, it is necessary to re-use software developed with different technologies, documented in different ways and list but not last hard to connect each other: paradoxically the re-use of the software could introduce new complexity within an existing Information System, and consequently increase the so called stove pipes.
In the last two years, in order to stimulate the use of SDMX within the ESS, Eurostat has started the development of an SDMX infrastructure that countries can re-use to speed the implementation of SDMX as common standard for sharing data and metadata, and harmonizing statistics. Now that the “SDMX NSI Reference Infrastructure” is available and enough mature, the next step is to foster a community of developers from the National Statistical Institutes and other statistical Organizations with the main aim to encourage further developments based on that Infrastructure, so to provide to the SDMX community a set of software that can be re-used partially or as a whole, but that they can guarantee homogeneity and interoperability.
The WP has the objective to create, through the tasks described below, the necessary instruments that will allow the community of developers to easily contribute to the further development of the “SDMX NSI Reference Infrastructure”.

Description of work
This WP will take as input: (i) the deliverables of the ESSnet on SDMX phase I with particular attention at WP4 (Generalized SDMX Framework), (ii) others SDMX software currently used in NSIs and available as open source, (iii) dissemination database schemas not still handled by the “SDMX NSI Reference Infrastructure” (e.g statistical data warehouse based on OECD.STAT), (iv) the current available software and documentation of the “SDMX NSI Reference Infrastructure”.The work and the deliverables will take in account the version of SDMX in which tools have been developed.
In order to achieve the objective, the following activities will be performed:
• In-deep analysis of the software modules and documentation of the “SDMX NSI Reference Infrastructure”. It will be identify where the source code and the related manuals are not adequately documented, suggesting improvements. Furthermore a set of step-by-step tutorial with examples on how to improve/enhance the current software modules;
• Investigate other SDMX software currently used within statistical organizations, with the main scope to identify useful functionalities that could be integrated within the “SDMX NSI Reference Infrastructure”;  
• Analyze the feasibility to integrate within the “SDMX NSI Reference Infrastructure” some software modules developed during the ESSnet phase I WP3 and WP4 (e.g EGR and Reference Metadata) and produce the appropriate guidelines;
• Analyze other statistical database not still covered in the current version of the “SDMX NSI Reference Infrastructure” and produce the use-cases;
• Design adequate templates that could support the evolutive maintenance of the “SDMX NSI Reference Infrastructure” within the community of developers.

Sub-Tasks:
T 4.1 Analyze the “SDMX NSI Reference Infrastructure”, its software module and its documentation. This work can require some added travel in Greece to the society that had developed the infrastructure. We have foresseen three travels: WP4-1, WP4-2, WP4-3.
T 4.2 Investigate other SDMX software currently used within statistical Organizations and Analyze their functionalities to be integrated within the “SDMX NSI Reference Infrastructure”.
T 4.3 Analyze the best way on how to incorporate, within the “SDMX NSI Reference Infrastructure”, some software modules developed during the ESSnet phase I WP3 and WP4.
T 4.4 Analyze other statistical database not still covered in the current version of the “SDMX NSI Reference Infrastructure”.
T 4.5 Design a set of templates that could support the evolutionary maintenance of the “SDMX NSI Reference Infrastructure” within the community of developers.

Deliverables (brief description and month of delivery)

D 4.1 (M16)
(a) Production of a technical document on “where to improve the documentation of the current software modules and manuals”
(b) Production of a step-by-step tutorial on “how to enhance/improve the SDMX NSI Reference Infrastructure” with source code examples

D 4.2 (M12) Production of a technical document with the appropriate use-case describing functionalities already existing in other SDMX software that can be integrated within the “SDMX NSI Reference Infrastructure”

D 4.3 (M8) Production of a technical analysis on how to incorporate modules developed in the ESSnet phase I WP4 into the “SDMX NSI Reference Infrastructure”: (a) EGR modules (b) Reference Metadata modules (c) SDMX Data Express

D 4.4 (M4) Production of a technical document describing the characteristics of other statistical database not still covered in the current version of the “SDMX NSI Reference Infrastructure” that can be incorporated in it

D 4.5 (M12) Production of a document contains the specification of a set of templates to be used for further developments of the “SDMX NSI Reference Infrastructure” by the community of developers

WP4 - Harmonization of the existing SDMX tools and software

<table>
<thead>
<tr>
<th>Main Tasks / Time schedule</th>
<th>Time Schedule</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M1</td>
</tr>
<tr>
<td>Analysis of “SDMX NSI Reference Infrastructure”</td>
<td></td>
</tr>
<tr>
<td>Investigate other SDMX software</td>
<td></td>
</tr>
<tr>
<td>Analysis of integration of ESSnet on SDMX software developed</td>
<td></td>
</tr>
<tr>
<td>Analysis of other statistical databases schemas</td>
<td></td>
</tr>
<tr>
<td>Design of a set of templates for documentation</td>
<td></td>
</tr>
</tbody>
</table>

Task/Work package number 5  Start Month or starting event: M1
End Month M18

Title Capacity building actions and dissemination of the ESSnet results

Participant id 1 IT coord 2 3 4 5 6 7

Objectives
Several capacity building actions are already put in place by Eurostat and other Organizations. In particular Eurostat within the ESS has been providing SDMX training sessions for Statisticians and IT Staff. Moreover a set of self learning tutorials (composed by videos, student books and self-tests) are available for downloading from the SDMX website. Finally the SDMX User Guide was updated in 2009.

Agreement number 60501.2010.001-2010.685
The idea of this WP is to not overlap the above actions, but transfer knowledge and experiences from NSIs to NSIs. If the version of SDMX will be released, the work and the deliverables will take in account the SDMX standard 2.1.

### Description of work

This WP will take as input: (i) the available training documentation, the self-learning tutorials, and the SDMX User Guide; (ii) the programs of the current available SDMX Trainings provided by Eurostat directly or indirectly through ESTP; (iii) the presentations provided during the bilateral meetings between Eurostat and the MSs; (v) the SDMX workshops that took place in 2009 at UNECE (in Geneva) and at UNESWA (in Beirut). Those two workshops were based on the agenda suggested by the SDMX Secretariat and oriented to management, statistician an IT staff.

In order to achieve the objective, the following activities will be performed:

- Analysis of the available training documentations
- Analysis of the technical documentation of the SDMX NSI Reference Architecture
- Definition of the agenda of a workshop with the main aim to transfer knowledge from NSIs to NSIs, organized on the model of UNESWA experience in 2009 (duration two days: ½ day plenary, 1 day parallel sessions between statisticians and IT staff, ½ day plenary)
- Produce a workshop documentation (slides and guides).
- Presentation to:
  - WP5-1 Global SDMX conference (2011)
  - WP5-2 MSIS (2012)
  - WP5-3 METIS (2012)
  - WP5-4 STNE (2011)
  - WP5-5 STNE (2012)
  - WP5-6 Metadata Working Group 2011
  - WP5-7 ITDG (2011)

### Sub-Tasks:

- **T 5.1** Analysis of the available training documentations and technical documentation of the SDMX NSI Reference Architecture (M2)
- **T 5.2** Design a questionnaire to be sent to countries in order to analyze the best way to define, organize and provide the capacity building actions, taking in consideration the difficulties in approaching and using the SDMX standard (M1)
- **T 5.3** Production of the workshop documentation. This documentation will be produced with in mind the reusability. In fact other workshops, seminars and trainings could be based on this documentation (M7)
- **T 5.4** Production of an handbook on how to use SDMX in statistical organization from a statistician point of you (M9)
- **T 5.5** Workshop on SDMX. The workshop will be hosted in Italy or in another available country and it will be opened to all countries that want to participate. This work package includes in its budget travel and daily allowance costs for each of 27 countries of the European Union, for two participants (one statistician and one IT staff).
  
  The workshop will be organized in two days as follow: ½ day plenary session with both statisticians and IT staff; 1 day parallel session for statisticians and IT staff; ½ day plenary session

- **T 5.6** Dissemination of the ESSnet results through a web site

### Deliverables

- **D 5.2** Questionnaire to be sent to countries and document with an analysis of the results
- **D 5.3** Workshop documentation: agenda, slides, workshop user guide
- **D 5.4** Handbook on SDMX.
- **D 5.5** Workshop on SDMX.
- **D 5.6** Dissemination of the ESSnet results through a web site
1.5. Distribution of manpower for the duration of the project

<table>
<thead>
<tr>
<th>Co-ordinator</th>
<th>WP1</th>
<th>WP2</th>
<th>WP3</th>
<th>WP4</th>
<th>WP5</th>
<th>TOTAL per Beneficiary</th>
</tr>
</thead>
<tbody>
<tr>
<td>(IT)</td>
<td>11,2</td>
<td></td>
<td>16,7</td>
<td>4,5</td>
<td></td>
<td>32,4</td>
</tr>
<tr>
<td>(SE)</td>
<td></td>
<td>14,5</td>
<td></td>
<td></td>
<td></td>
<td>14,5</td>
</tr>
<tr>
<td>(SK)</td>
<td></td>
<td></td>
<td>49,5</td>
<td></td>
<td></td>
<td>49,5</td>
</tr>
<tr>
<td>(BG)</td>
<td></td>
<td></td>
<td></td>
<td>30,3</td>
<td></td>
<td>30,3</td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>126,7</td>
</tr>
</tbody>
</table>

1.6. Dissemination towards the ESS

The Dissemination is part of the WP5. The web site used in the ESSnet of SDMX first phase (Eurostat Wiki) can be used also in “ESSnet on SDMX II” to made available information about project, deliverables and monthly reports. The site is planned to have a wiki corner to stimulate the debate related with the topics among the users within and outside the ESS. The presentation of intermediate and final results will be made in different meetings and working groups: Global SDMX conference, MSIS, METIS, STNE, Metadata Working Group, ITDG.

1.7. Management

<table>
<thead>
<tr>
<th>Task/Work package number</th>
<th>1</th>
<th>Start Month or starting event:</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>End Month</td>
<td></td>
<td>X/End of project</td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>Project management</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant id</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>IT</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Objectives

Description of work

The co-ordinator country will:

- Ensure that the Action produces the required deliverables accordingly to the identified standards of quality, on time and within the budget;
- Maintain continuous contacts with Eurostat and partner countries;
- Follow-up of the different Work Packages (WP’s) and produces quarterly progress reports;
- Define progress reports structure and deliverables reports structure;
- Produce exceptional alert reports if unexpected events affecting seriously the project occur;
- Identify changes to the plan and submit them to the Steering Group for approval;
- Identify problems and risks and propose solutions to overcome them;
- Make deliverables available to the Steering Group for evaluation and approval;
- Gather cost statements and provide payments to partner countries accordingly with payment rules and the work effectively done;
- Monitor the updating of the ESSnet on SDMX site with all the WP’s documentation and deliverables;
• Promote, organize and manage the meetings planned. This work package includes in its budget travel and daily allowance costs for all the partners for the planned meetings (1 Kick-off, 4 coordinator-meetings)

**Sub-Tasks:**

1. Checks and jointly approves with Eurostat each WP definition, schedule and budget
2. Organization and coordination of the partners
   - Definition of quarterly report structure
   - Definition of deliverable report structure
3. Organization of the management and control meetings
4. Monitoring of the different WP's
5. Accept the contractual deliverables on a provisional basis;
   - Production of progress quarterly reports of the Action
6. Asks each WP responsible a monthly report
7. Keep a centralized register of changes (scope, timescale, cost)
8. Keep a centralized register of the risks (with mitigation plans)
9. Management of all information flow with Eurostat/Steering Group
10. Gathering cost statements;
11. Make payments to the partners accordingly with payment rules and the work effectively done;
12. Makes available to all participants collaboration tools, namely Circa Interest Group, and promotes the use of video and teleconference.

**Deliverables**

- Standard structure for the different kinds of reports to be produced along the action;
- Centralized registers of changes and risks;
- Quarterly progress reports (beginning on the end of M3)
  WP1-1 Kick-off meeting
  WP1-2 first coordinator meeting
  WP1-3 second coordinator meeting
  WP1-4 third coordinator meeting
  WP1-5 final report
- Exceptional alert reports;
- Costs and payments balance sheet;
- Pre-acceptance of final deliverables;
- Collaboration tools;
1.8. Adequacy to specific ESSnet criteria 5, 6 and 7

1.8.1. Contribution to knowledge sharing, innovation and harmonisation and plans for the use and dissemination to the ESS

The ESSnet phase II will reuse in large part the results obtained in the ESSnet on SDMX phase I. Below is described for every WP how this will be obtained:

WP2: From the description of the process for international reporting developed in Phase I, the work in Phase 2 will be focused to clarify the possibilities to also deal with more complex statistics in the developed process. This implies trials with the complex materials including the Census. Make possibly necessary developments in the input systems and make documentations and advice on how to use the concept for Statistics Sweden as well as the PC-Axis community. The work in Phase 2 will both include what is developed at Statistics Sweden as well at Eurostat in the SDMX Reference Architecture.

In Phase 2 also software developed in Phase I will be inserted in the Internet version of PC-Axis called PX-Web.

WP3: In the Phase-I we have concentrated on definition of e-questionnaire design, covering basically the Process 2 (Design) of GSPD model, its’ content and structure with the usage of metadata structure definition. We used standard SDMX tools (Metadata editor) to create a metadata messages, which can describe any questionnaire. In the Phase II, we following the GSBP model will continue further in the statistical business processes and create tools, which would automate the process of possible questionnaire data (or potential any microdata) dissemination. This means that we shall continue with using questionnaire metadata we collected in first phase for automatic generation of DSD for data dissemination, automatic mapping between DSD and physical tables in database and prepare data for dissemination.

WP 4: In sub task T.3 of WP4 will be used the results obtained from the WP3 and WP4 of ESSnet phase I:
- SDMX Data Express
- SDMX Istat framework

The main objective is to analyze the possibility to reuse code partially or completely in order to integrate the NSI Eurostat Reference Architecture with new functionality and produce a suitable analysis document.

In fact the issue arose in the last years by the NSIs was the possibility to have not several framework in supporting exchange of data in SDMX but only one that could satisfy all the different needs.

WP5

The WP5 is focused on the capacity building, therefore it is not foreseen to reuse any results of ESSnet phase I.

---

1 For ESSnet criteria see document CPS 2007/61/6 - 61st Statistical Programme Committee of 15 February 2007. The criteria are: 1. Involving several partners and results diffused to non participating NSAs; 2. Focus on issues that answer to a European interest; 3. Compatible with the 5 year programme; 4. Cost effective; 5. Knowledge sharing, innovation and harmonisation; 6. Sustainable; 7. Actions must be carried out by ESS organisations.
The results of work will be developed in the ESSnet on SDMX phase II will take in consideration the first indicator for evaluation of the action “The quality of the results and deliverables achieved, in particular with regard to their reusability within the European Statistical System”. The results represented by tools and documentation will make SDMX more comprehensible both from IT staff and Statistician so to make more easy the implementation of SDMX by countries. Below we will indicate for every WP the level of reusability of the work for other organizationi in ESS:

- WP2: All the work in Phase 2 is targeted to the reusability for other NSI:s. More integrated in the PC-Axis software and more use cases and more documentation. Also workshops are planned for NSI:s in the PC-Axis community on the first hand but other NSI:s can also participate as observers to assess if they can learn from our results or perhaps join the PC-Axis community.

- WP3: So far in the NSI there are not experiences with micro data collection using SDMX standard and its eventual dissemination. While developing the SDMX supporting tools in the WP3 of the proposed project, very important experience concerning dissemination of Microdata could be gathered. These experiences would be a breakthrough in Microdata SDMX exchange. They can be than used whenever a need for Microdata exchange raises. The methodology, deliverables and tools developed could after adaptation to the national specifics be used by NSIs in the data collection and/or micro data dissemination.

- WP4: is oriented to foster a community of developers from the National Statistical Institutes and other statistical Organizations with the main aim to encourage further developments based on Eurostat SDMX infrastructure, so to provide to the SDMX community a set of software that can be re-used partially or as a whole, but that they can guarantee homogeneity and interoperability. This will be reached through an in deep analysis of the design architecture and the source code of the different building blocks that are part of the Eurostat SDMX NSI Reference Infrastructure. The analysis will identify where and how the user can intervene to adapt the software. The interaction with Eurostat and its contractors should considered to the maximum extent in order to coordinate the actions foreseen in this WP with the enhancements that Eurostat has already planned in 2011.

- WP5: This WP will deal with mainly with a set of Capacity building actions. In this context the main aim is to increase the knowledge and comprehension of SDMX with particular regard to Statisticians. All the deliverables are thought with in mind the “portability”. In fact it will be designed a workshop and related documentation (agenda, slides and user guide) that can be reused not only for other events, but also as in regional seminars, and bilateral or multilateral meetings. The workshop and the related documentation will be based on the idea to stimulate a discussion between IT staff and statisticians with the main objective to facilitate the introduction of SDMX within the National Statistical Institutes.. The handbook will be designed with in mind to develop a reusable document that is more explicative of a leaflet, but less verbose of a self-learning tutorial and, in any case, the NSI point of view will be the main stream.

1.8.2. Sustainability

The ESSnet on SDMX phase II output will be support for NSIs on different aspects:

- use of SDMX standard at the NSI level in the phase of statistical micro-data processing ESS partnership

- adaptation of NSI reference architecture realized by Eurostat to their own needs

- using SDMX in pc-axis community

Agreement number 60501.2010.001-2010.685
• better understanding of SDMX by users and in particular statisticians

1.8.3. *ESS partnership*

The ESSnet on SDMX II group is composed by four countries will bring in the project their extensive knowledge and long term experience in SDMX and in statistical processes.