# SMS-Quality: the project and application focused on metadata on quality

Jitka Prokop, Czech Statistical Office, jitka.prokop@czso.cz

Key words: statistical quality, reporting, metadata

The paper **introduces the SMS-Quality project and application** focused on statistical quality metadata. It covers the **aims**, **general user structure** and **functionality** of the application, course of development and experience up to now.

The application is designed as a flexible tool for **monitoring** and **comparisons**, quantitative and qualitative **assessment** of product and process quality. It enables to design, compile and generate in ordinary user file-formats various **reports** in Czech and English, with either open or limited access. Via common interface it compiles data about individual phases of statistical process, with focus on data processing, user requirements, or dissemination; using metadata from SMS-Classifications. It looks at data on several levels, at individual metadata and their aggregations over reference periods or across various statistics etc. The metadata (numerical, textual or datum) can be **benchmarked**: either automatically or ad-hoc by an expert, usually ex-post, primarily for internal purposes.

The project was originally aimed to provide support for ESS quality reporting and for internal management, including self-assessment and auditing. It was based mainly on ESS quality indicators and criteria (dimensions) and DESAP checklist. Possible incorporation of ESS SIMS items and use of GSBPM are to be investigated.

## Statistical Metadata System on Quality: Introduction

### Aims, Functions, Focus

The aim of this paper is to introduce the concept and the software tool with ‘web-browser’ interface SMS-QUALITY. The development of this application has been a part of the total redesign of the Statistical Information System (SIS) and the Statistical Meta-information System (SMS) in the Czech Statistical Office. In the time of writing this paper, after analysis and improvements, the project is in the phase of the main programming together with ‘client’ (i.e. ‘user’) testing of ‘beta-versions’, before main debugging and full implementation. This chapter provides overview before describing some aspects more into details.

The SMS-QUALITY is focused on quality monitoring, reporting and self-assessment; and is also relevant for the use within auditing processes. The intention is to support improvement of quality reporting and statistical quality itself, by providing quality management with a relatively comfortable and flexible tool integrated within the Statistical Information System. The application also enables benchmarking designed and considered primarilyforinternalmanagement purposes. Using cross-cutting overviews’ options, results from one or more various surveys can be compared or aggregated and assessed, taking into account user needs and demands, priorities and trade-offs.

The main focus in the application starts with quality of a particular statistical survey, and then it can lead to comparisons, aggregations and assessments for a group of surveys. By a survey we mean any statistics (or a survey) that runs or at least has data stored in the central database. It covers business, social and demography statistics, partly also national accounts, price statistics or administrative data statistics.

Within the application, different types of reports can be defined. Historically, it was inspired by the standard ESS quality reports (producer-oriented) [1,2], by quality criteria and quality (performance) indicators [3], and by the structures like the DESAP checklist [4]. Proposed structure is a combination of quality criteria and the phases of statistical process (referring to GSBPM, [5]). Other structures can be defined as well, for example relating to the Code of Practice [6] agenda or EFQM. However, the main use is expected for the standard producer-oriented ESS quality reporting. The updated Single Integrated Metadata Structure (SIMS) [7] would be applicable; however the present project does not cover the export or conversion of outputs into SDMX format. Generally, SMS-QUALITY covers more information, compared with SIMS. It is generally designed for internal use, however, via a special parameter; some of the outputs are recognized as 'public' and available in ordinary data formats suitable for web-publishing.

### Design of Q-Maps, retrieval of values, and flexibility

The basic item in each report is what we call **Q-attribute** (or a quality item, indicator, measure). The Q-attributes are thematically focused on quality in a broader sense, and they **cover** (though not always exhaustively) various stages of statistical process, dissemination, requirements of users, and the product quality (meaning both output data and publications). They can be either textual, numerical, or a date (for more details, see the chapter 2 below).

In the application, any design and preparation of a new report starts by the definition of the **structure** of these Q-attributes, i.e. *sections,* *sub-sections* and in them *Q-attributes[[1]](#footnote-1)* themselves. This structure is called a General Q-Map; and it also includes preliminary settings of parameters of the Q-attributes. For similar surveys (e.g. short term statistics), and in the next step the purposes of one concrete statistics (a particular survey), some specifications and adjustments are done by competent managers in the derived Specific and Survey Q-Maps.

Afterwards, according to previous definitions and specifications, values are generated from databases or inputted into a Value Q-Map. To enable **retrieval** of values that are already available, there are defined (individually for each Q-attribute) SQL queries into either the central database or the databases of other SMS applications. Quality indicators like coefficient of variation or (non)-response rate are stored as attributes of particular aggregated data in the central data warehouse; and shall be retrieved from there. Application also enables simple **calculations** (e.g. for quality criteria such as punctuality) using the ‘collected’ data. Other values (e.g. textual evaluations etc.) are **inputted** into the application directly, **manually**, by a user.

Before the final **approval** of any concrete Value Q-Map, all the values (i.e. those retrieved from a database, calculated or directly inputted) can be **edited**. This does not change the source data, because the application uses a separate database within the IT system and stores both original and modified values there.

The process ends by editing and a final approval of the whole Value Q-Map by a survey manager. Due to the suggested way, content of any report is quite **flexible**; and ‘the only’ main thing that is generally fixed is the number of hierarchical levels of Q-Maps and suggested parameters.

### Q-Maps

As mentioned above, there are several levels of so called Q-Maps which need to be specified for each report (e.g. ESS quality report). The hierarchy of Q-Maps is the following: 1) **General Q-Map** (**GQM**) is a base for all other Q-Maps. The number of general Q-Maps is not limited. 2) **Specified Q-Map** (**SpQM**) is designed for a group of similar surveys (generally statistics), where the requirements for reporting are similar; and comparisons and aggregations are planned between surveys. Q-attributes can be added only into the general Q-Map; and in the specific Q-Map are either confirmed for use, or cancelled. 3) **Survey Q-Map** (**SuQM**) is derived from a Specific Q-Map and adjusted to a concrete statistical survey: a manager or a statistician specifies the (key) statistical variables to enable evaluation of their results; and he can adjust some benchmark parameters (see below). 4) While the Q-Maps mentioned above are in fact the defined structures and parameters of Q-attributes, a **Value Q-Map** contains real values of the Q-attributes and benchmark-results for a concrete statistical survey, data processing[[2]](#footnote-2) and reference period.

**Compatibility** among these maps is kept by the application and by the manager; the application is not fully dynamic. When a new version of a relevant Q-Map is approved, managers (owners) of derived Q-Maps shall be informed by email; and either **update** their Q-Map or keep the version before the update.

### Q-Forms: comparisons, aggregations and self-assessment

One **Value Q-Map** covers only one reference period. To provide overviews over several reference periods, statistical variables or surveys, there is a tool which we call the **Q-Form**. It is designed for comparisons, aggregations, self-assessment and also as a supportive tool for auditing. Besides the same options used in the above mentioned Q-Maps, the Q-Forms use data and results of benchmarking from several Value Q-Maps.

### Connection of SMS-QUALITY with other SMS applications and databases

The SMS-QUALITY application is **interlinked** with other SMS applications (subsystems). It uses particular nomenclatures[[3]](#footnote-3) from SMS-CLASS, which is the only source for definition, maintenance and update of nomenclatures or classifications (besides internal application codings). There are several applications or subsystems, which shall supply data to SMS-QUALITY. The information related to *accuracy, timeliness and punctuality of processes, or comparability* shall be gained via SMS-SURVEYS[[4]](#footnote-4), generally focused on statistical processes. The information on statistical products and the service for users, including data to *relevance*, *accessibility*, *clarity*, *timeliness* and *punctuality* of outputs will be gained from SMS-DISSEMINATION. The SMS-REQUIREMENTS application manages information regarding user needs (namely concrete requirements) in relation to concrete surveys (statistics), so data regarding to *relevance* will be retrieved from there. The SMS-TIME-SERIES application manages time series of aggregated data. The survey manager in SMS-QUALITY defines which concrete time-series shall be evaluated by applying relevant quality measures (indicators, Q-attributes).

## Quality Attributes (Items, Measures, Indicators)

This chapter introduces aspects of Q-attributes into more details. As mentioned above, a Q-attribute is the basic item of any Q-Map. In the application, it depends on a user which Q-attributes he chooses into her Q-Map or Q-Form. If it is meaningful, values of Q-attributes can be benchmarked.

### Q-attributes as qualitative and quantitative metadata

Providing information on quality in broader sense, Q-attributes are practically **metadata** and they might be either quantitative or qualitative. For 1) **quantitative** indicators, the ESS quality performance indicators (like response rate or coefficient of variation), or other relevant ones are considered. For 2) **qualitative** indicators, more types have been indicated: a) *reference metadata*, i.e. information about some aspect(s) of a survey which can be supportive for further assessment of quality e.g. within self-assessment or auditing; b) indicators related to *process*, and possibly process management as well, it includes improvement actions, etc.; and c) expert *evaluation of* *results*.

### Methodological categories of Q-attributes

For better recognition, **categories of Q-attributes** have been methodologically indicated and proposed; to improve orientation during methodological design work, searches and filtering. Up to now, there are the following categories defined: 1) *Basic information* (referring to a statistical survey); 2) *user requirements* and their handling, it covers reasons for a refusal or acceptation, and monitoring of implementation; 3) *methodology*, namely *imputations*, *small area estimations*, *estimations*, *variance*, *seasonal adjustment*, *and sampling*; 4) *time schedules* of a survey; it includes indicators on planned and real dates, *timeliness and punctuality* for concrete activities or actions; 5) *statistical* *process phases*: design, preparation, data collection, (central) processing; 6) *data confidentiality and data protection*; 7) data sources and files, i.e. *frame* and *sample*, and *secondary data sources* (e.g. administrative data, registers, other surveys); 8) *outputs and dissemination, information service*; 9) *quality dimensions (criteria)* covering *relevance, accuracy, accessibility, timeliness and punctuality, accessibility, comparability,* *coherence*; and 10) *quality indicators* (quantitative)*.*

### Levels of Q-attributes

In the beginning of the project, three levels of Q-attributes were indicated (and described in the papers [8, 9]), depending on general stability of their values and the level of their detail. These three ones were updated into the following five levels; one Q-attribute in the application can have defined only one level: 1) **General** (or Super-Task) **level**. For this level, any special functionality has not been indicated; however it is available for further development, and for aggregations. 2) **Statistical survey** **level** includes the Q-attributes which directly relate to a statistical survey, so the values generally do not necessarily change with new processing or year. Examples of Q-attributes on this level are e.g. key users, methodology etc. 3) **Reference year** is newly added level and is independent on periodicity of statistics. It might be useful for annual quality reporting. 4) **Processing level** relates to all reference periods (incl. revisions) processed in one time. In the end it usually covers only approved values, produced by the statistics. Q-attributes on this level are e.g. unit response rate, sample size etc. 5) **Reference period** level refers not-surprisingly to an individual reference period. Individual Value Q-Maps are exported separately for each reference period, compared with Q-Forms. Q-attributes on this level are e.g. coefficient of variation, item response rate etc.

The level of Q-attribute is related to the **functionality** of the application. If a new reference year, a new processing or a reference period is established, **validity** of the values on relevant level is ended automatically. Some **user roles** are related to the reference year, for example an editor of Value Q-Maps or a viewer of not yet approved Q-maps. To **export** values, a user needs to choose concrete reference period or can export more reference periods at one time.

### Parameters of Q-attributes

For Q-attributes, various *parameters* have been suggested. The values of these parameters are already pre-specified in the nomenclature of Q-attributes; however they can be revised and modified in the SMS-QUALITY application. Some of them are relevant only for selected Q-attributes. The parameters are the following:

1) **Category** refers to methodological content (see categories of Q-attributes); 2) **Level** refers to the level of detail and validity in time. 3) **Benchmark** marks potential suitability for categorical assessment (i.e. benchmarking). 4) **Source of Values**refers to SMS-subsystem from which the values are to be retrieved and **Nomenclature** specifies the nomenclature from which the values of some Q-attributes are to be selected. 5) There are also more technical parameters, concretely a) **Data Format** (meaning text, long text, number or date; or a textual code from nomenclature), b) **Data Mask**; and c) **Multiplicity** **of Values** which determines the (maximum) number of values, which can be valid and showed the same time. For hierarchical relations between Q-attributes, d) the parameter **Parent-Child** is available. For the child Q-attribute is specified its parent and e) the **Layout** how to display results (transposition, tab, or tree).

## ‘Q-Forms’ – Quality Maps for comparisons and aggregations

As mentioned in the introduction, Q-Forms are primarily designed for **comparisons**, **aggregations** and **assessment** of values and benchmark results of Q-attributes from more reference periods, years, statistical variables (meaningfully their benchmark results), and surveys.

The **hierarchy** of Q-Forms is simpler compared to Q-Maps; however the process of their definition and functionality are similar. Generally, there are General Q-Forms, Specific Q-Forms derived from them, and Value Q-Forms generated subsequently. A **General Q-Form** is a specific type (namely a parameter) of the General Q-Map. In the **Specified Q-Form** the aggregations are defined; and the values are then generated into a **Value Q-Form**, and edited there.

More concretely, within the Specific Q-Form, the user chooses concrete Value Q-Maps (within one General Q-Map) and specifies all the dimensions (surveys, years, statistical variables, benchmarks, releases and reference periods) individually for each Q-attribute in a Q-Form. He can also define aggregations of values or benchmark results from Value Q-Maps. There are available basic functions like sum or average; and how to treat null values has to be decided.

After the definitions are ready, available values can be retrieved on ‘one-click’ from source databases into derived Value Q-Form. Before the approval by a manager, the same as for Value Q-Maps, all the values can be edited (i.e. changed or imputed) by an assigned editor. Both original and changed values are archived and can be shown; the change is visually marked; and a comment can be added to any result (value).

## Benchmarking

Benchmarking is designed and intended primarilyfor **internal management purposes**. It is categorical assessment where the output value is assigned to one of pre-defined categories. Preferably, the benchmarked values are either numerical or textual; incl. values from nomenclatures. For the ‘date’ type of data, calculation of timeliness or punctuality is suggested before benchmarking.

In the application, the decision on benchmarking is to be done separately for each Q-attribute. The following **parameters** have to be specified:

1. *‘Shall the values of this item be benchmarked’*? This is the first, ‘existential’ decision.

2. *The way of benchmarking*. The benchmarked values might be assigned to a concrete category either individually in each case (by an expert; it is useful especially for textual values), or automatically, according to pre-defined parameters.

2. *Categories*. The application suggests up to **7** categories for benchmarking and it is up to the manager how many of categories will be used. Except for manual benchmarking, the manager needs to pre-define either boundaries, or to assign generally each item from the used nomenclature to a particular category. If relevant, one category can be defined by one or more intervals, e.g. by two ‘opposite’ intervals {<-3;-2); (2; 3>}.

3. *Description of categories (labels)*. During the design of a concrete Q-Map, categories need to be labelled (i.e. described by a word): either directly in the application, or using a nomenclature with already pre-defined sets of ‘labels’.

Referring to the levels of Q-Maps and the user roles, definitions and adjustments are allowed only on certain levels. The benchmarking suggested and defined in a General Q-Map can be ‘switched off’ or accepted for a use on the level of Specific Q-Map (separately for each Q-attribute). The number of categories is decided and fixed in the General Q-Map; the way of benchmarking (automatic or manual) on the level of Specific Q-Map. The interval boundaries or the ‘coded values’ assignment can be adjusted in related Specific and Survey Q-Maps.

This approach enables to compare results of various surveys, taking into account user requirements, priorities and trade-offs. For example, demands regarding accuracy (namely coefficient of variation or response rate) might be different for short term statistics and annual business surveys respectively (or even for various strata within a survey). However because the way of benchmarking stays unified at least for similar statistics (on the level of a concrete Specific Q-Map), aggregations or comparisons between different statistics are (also technically) possible.

## User roles, accessibility and availability to public

The application SMS-QUALITY is generally designed for internal use, especially for producers of concrete statistics (survey managers, statisticians) and internal managers, but the provision of comfortable outputs to any users have been taken into account as well.

During the analysis, different types of users have been defined; and in relation with that the parameters of Q-Maps and Q-Forms were suggested. Either an individual Q-Map or just a selection of Q-attributes in this Q-Map (mainly a Survey Q-Map) can be marked as ‘available to public'. Similar parameterization enables the recognition of non-approved Q-Maps (esp. Value Q-Maps) that are available only to limited number of users. Approved and non-public ones are available to all internal users.

Regarding the user roles, of cause, for each concrete Q-Map or Q-Form one ‘owner’ is assigned, that can define editors and owners of lower level Q-Maps. Other users can have granted access as viewers.

## Output formats and languages

The outputs can be exported in various **formats**, namely PDF, HTML, XLS, DBF, DOC, RTF; and these files can be provided to users or put on website. SDMX format has not been considered during the planning and design of this project, partly because of timing, incl. of the related agenda. Otherwise, SMS-QUALITY could be considered as a suitable tool for the preparation of ESMS quality reports. The solution for presentation of outputs in two languages (here Czech and English) is also in a process of design and programming.

## Conclusions

SMS-QUALITY enables cross-cutting compilations, monitoring, comparisons, aggregations, evaluations and self-assessment. It is designed as a supportive tool for management and covers both statistical process and product quality; using as much as possible data and meta-data already available in the data warehouse. One of its strength is the flexibility of content and settings that are manageable directly by managers in the application, without any need of programmer knowledge. The application is suitable for a preparation and compilations of the ESS reports in SIMS structures; however the direct use for this purpose will depend whether outputs in the SDMX format are designed in any follow-up project in the future. At the present, the project enables outputs in ordinary formats, including in English, and distinguishes public outputs from internal ones. The data are generated and the outputs exported upon the request of assigned users.

## References

[1] Eurostat (2009), [ESS Standard for Quality Reports](http://epp.eurostat.ec.europa.eu/portal/page/portal/lang-en/ver-1/quality/documents/ESQR_FINAL.pdf), ISBN 978-92-79-07854-5.

[2] Eurostat (2009), [ESS Handbook for Quality Reports](http://epp.eurostat.ec.europa.eu/portal/page/portal/lang-en/ver-1/quality/documents/EHQR_FINAL.pdf), ISBN 978-92-79-07855-2.

[3] Eurostat (2013), [ESS Guidelines for the Implementation of the ESS Quality and Performance Indicators (QPI)](http://epp.eurostat.ec.europa.eu/portal/page/portal/quality/documents/Draft_template_quality_performance_indicators_2013.pdf), Doc. ESTAT / B1/AB D(2013).

[4] Eurostat, [The European Self Assessment Checklist for Survey Managers (DESAP)](http://epp.eurostat.ec.europa.eu/portal/page/portal/quality/documents/desap%20G0-LEG-20031010-EN.pdf), http://epp.eurostat.ec.europa.eu/portal/page/portal/quality/documents/desap%20G0-LEG-20031010-EN.pdf

[5] UNECE (2013), [Generic Statistical Business Process Model. Version 5.0.](http://www1.unece.org/stat/platform/display/GSBPM/GSBPM+v5.0)

[6] Eurostat (2011), [European Statistics Code of Practice](http://epp.eurostat.ec.europa.eu/portal/page/portal/product_details/publication?p_product_code=KS-32-11-955), ISBN: 978-92-79-21679-4.

[7] Eurostat (2013), Technical Manual of the Single Integrated Metadata Structure (SIMS), Doc. ESTAT/B1/AB D(2013).

[8] Prokop J., Melíšková J., Petrikovits E., Czech Statistical Office, (2010), [The Quality Metadata System in the Czech Statistical Office](http://www.unece.org/stats/documents/2010.03.metis.html), Joint Work Session on Statistical Metadata (METIS), UNECE/Eurostat/OECD, 10-12 March 2010, Geneva, Switzerland.

[9] Prokop, J. (2010), [The CZSO Quality Metadata System and It’s Use in Quality Monitoring, Assessment and Methodology Auditing](http://q2010.stat.fi/papersbig/). European Conference on Quality in Official Statistics, 4-6 May 2010, Helsinki, Finland.

1. Earlier called a *section group,* a *section* and a *Q-attribute*; see [8, 9]. [↑](#footnote-ref-1)
2. Data processing means here ‘a batch of processes’ running within one time period; besides first estimates it might include revisions. [↑](#footnote-ref-2)
3. By a ‘nomenclature’ meaning a one-level classification (i.e. without any hierarchy). [↑](#footnote-ref-3)
4. Originally called SMS-TASKS; by a ‘task’ meaning a statistical survey of any type. [↑](#footnote-ref-4)