Workshop for Developing Countries on the Revision of the International Recommendations for International Merchandise Trade Statistics

Bangkok, 9-12 September, 2008

Data Quality and Metadata

(Providing guidance on data quality (reporting, assessment and indicators); Ensuring data quality; Compilation and dissemination of metadata; Main elements of metadata)

Presentation by Markie Muryawan
Statistician,
United Nations Statistics Division, International Trade Statistics Section,
2 United Nations Plaza, DC2-1540, New York, New York 10017,
Phone: (212) 963-0692, Fax: (917) 367-5106, e-mail: markie@un.org

Data Quality

1. Current recommendation(s)

IMTS, Rev.2 covers the issue of data quality mostly from the perspective of reporting and dissemination (see chapter VII). IMTS:CM contains a chapter on issues of data quality control (IMTS:CM, chapter 11).

2. Result from the 2006 NCDP Questionnaire

Question: Not available	**Yes**	**No**	**N/A**

Since the introduction of IMTS, Rev.2 in 1998, a number of quality frameworks, reports and indicators were developed and implemented by international and regional organizations (IMF, OECD and Eurostat). Furthermore, separate chapters on data quality were included in newly developed international recommendations in other areas such as distributive trade statistics and industrial statistics.

Therefore, it is suggested to cover the issue of data quality also in IMTS, Rev.3.

3. Result from world wide consultations

Question 14: Do you agree				
(i) that IMTS, Revision 3, should provide recommendations on quality of international merchandise trade statistics	**Yes**	**No**	**No Op.**	**N/A**
Total of 110 replies	91	1	6	2
Developed (34)	85	3	12	0
Developing and transitional (76)	93	0	4	3
(ii) that this includes reporting of data quality?	**Yes**	**No**	**No Op.**	**N/A**
Total of 110 replies	82	3	13	3
Developed (34)	82	6	12	0
Developing and transitional (76)	82	1	13	4
(iii) that this includes assessment of data quality?	**Yes**	**No**	**No Op.**	**N/A**
Total of 110 replies	82	4	13	2
Developed (34)	76	9	15	0
Developing and transitional (76)	84	1	12	3
(iv) that this includes the use of some data quality indicators?	**Yes**	**No**	**No Op.**	**N/A**
Total of 110 replies	87	2	10	1
Developed (34)	85	3	12	0
Developing and transitional (76)	88	1	9	1

International MerchandiseTrade Statistics Section, September 2008

Data Quality

4. UNSD observation of wwc result

There is almost universal support for all four parts of this proposal (91%, 82%, 82% and 87% respectively with no-answers for all parts of the question not exceeding 4%).

Despite this overall support numerous countries have strong concerns about the scope of these recommendations (for example whether indices and surveys or asymmetries will be covered) and how such guidance will be provided.

There seems to be a preference for guidelines or description of best practices/ examples to be provided in the Compilers Manual instead of recommendations.

Not surprisingly, mainly developed countries (in particular EU members but also the USA) want that such recommendations/ guidelines are not in conflict with their established practices.

Data Quality

5. Selected comments (from total of 74 comments)

- "We would like to see it stressed that quality must be conceived as a "strategy" and not as an optional aspect referring to some pieces of the production process."

- "The assessment of data quality can be extremely difficult. The inclusion of such a recommendation would require careful wording and instructions for this measurement."

- "When comparing different countries' trade data for economic monitoring and trade policy purposes, it is useful to have an assessment of quality of that data and especially an indication of the likely effect on asymmetries."

- "Another way to increase the quality is to calculate trade indices. While calculating, additional unit value checking is necessary. Sometimes, the error that we could not find during editing trade figures can be found by the staff who is responsible for trade indices."

Existing Data Quality Frameworks

IMF Data Quality Assessment Framework (DQAF)

- The IMF DQAF views quality through a prism that covers <u>governance of statistical systems</u>, <u>core statistical processes</u> and <u>observable features of the outputs</u>.
- The framework is applicable to countries in all stages of development as it comprehensively covers the various quality aspects in which data are collected, processed, and disseminated.

Eurostat Quality Definition

- Eurostat's basic aim is to define <u>quality of the output</u> of official EU statistics.
- The starting point is the user and how he/she sees the product, and Eurostat works its way back to the underlying processes where the outputs do not yield a direct measurement.
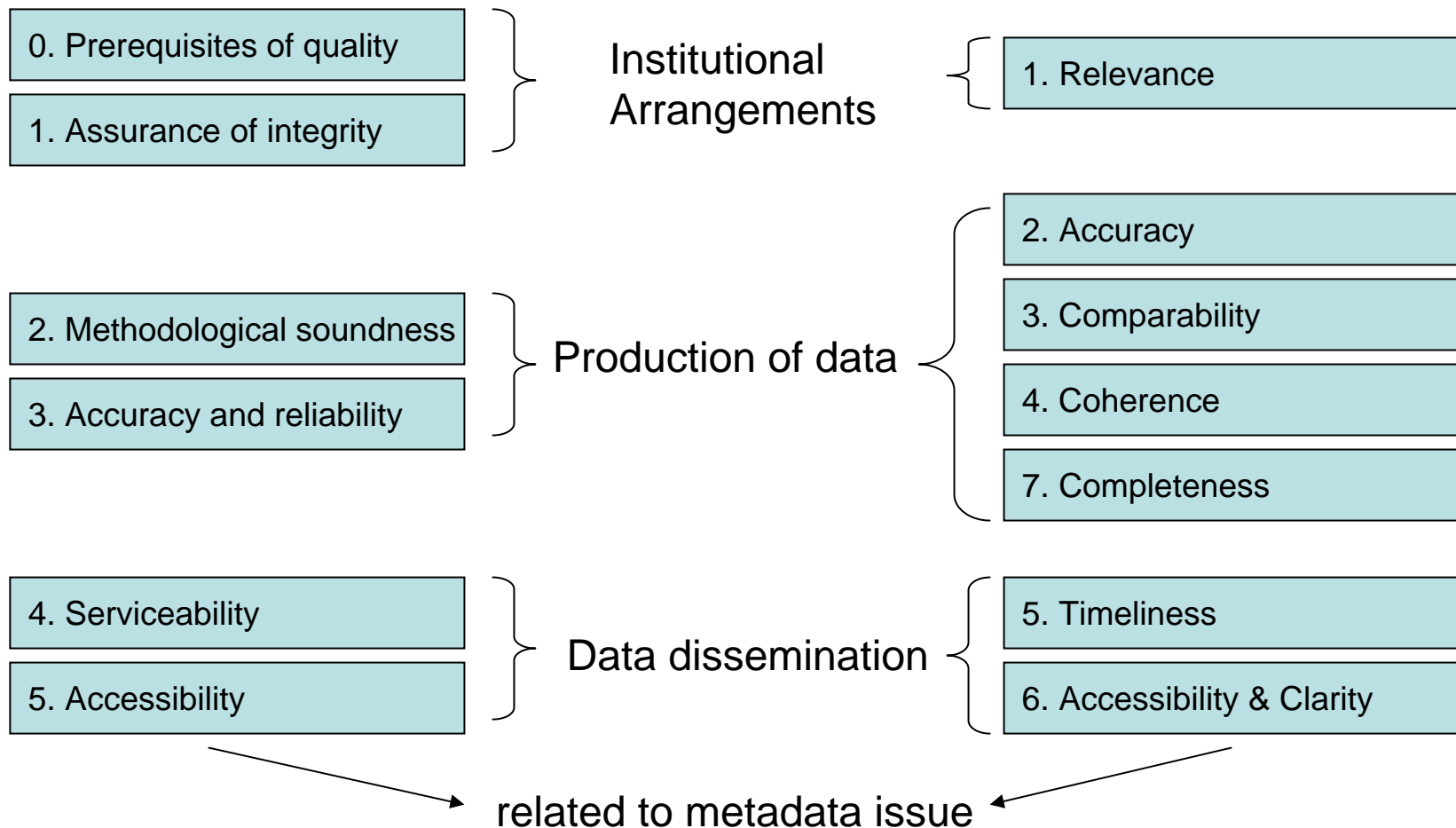
UN comtrade 🌐 | united nations statistics division

Quality Aspects Addressed
by IMF DQAF

Key Questions	IMF DQAF
How the quality of statistics is affected by the legal and institutional environment and resources, and is there quality awareness in managing activities?	0. Prerequisites of quality
What are the features that support firm adherence to objectivity in the production of statistics so as to maintain users' confidence?	1. Assurance of integrity
How do the current practices relate to the internationally agreed methodological practices for specific datasets?	2. Methodological soundness
Are the source data, statistical techniques, and supporting assessments and validation techniques, inclusive of revisions studies, adequate to portray the reality to be captured by specific datasets?	3. Accuracy and reliability
How are users' needs met in terms of timeliness of the statistical products, their frequency, consistency, and their revision cycle?	4. Serviceability
Are effective data and metadata easily available to data users, and is there assistance to users?	5. Accessibility

Quality Aspects Addressed
by Eurostat Quality Definition

Key Questions	Data Quality Dimension
Are the data what the user expects?	1. Relevance
Is the figure close to the exact or true values?	2. Accuracy
Are the data in all necessary respects comparable across countries or periods?	3. Comparability
Are the data can be used with/for other data? Ex. BOP, SNA, etc.	4. Coherence
Does the user get the data in time and according to pre-established dates?	5. Timeliness
Is the figure easily accessible and understandable?	6. Accessibility and clarity
Doest amount of available data match what was expected to be obtained?	7. Completeness

IMF DQAF vs Eurostat Quality Definition

| 0. Prerequisites of quality |
| 1. Assurance of integrity |

Institutional Arrangements

| 1. Relevance |

| 2. Methodological soundness |
| 3. Accuracy and reliability |

Production of data

| 2. Accuracy |
| 3. Comparability |
| 4. Coherence |
| 7. Completeness |

| 4. Serviceability |
| 5. Accessibility |

Data dissemination

| 5. Timeliness |
| 6. Accessibility & Clarity |

related to metadata issue

Metadata

1. Current recommendation(s)

… Therefore, it is recommended that data compilers:
- Publicly disseminate documentation on their sources and methods;
- Publicly announce scheduled release dates;
- Provide regular monthly reporting of data to the user community through publications and/or electronic media;
- Regularly revise data (when additional information is available), taking into due consideration user needs for reliable statistics.

Refer to IMTS, Rev.2, para. 154

2. Result from the 2006 NCDP Questionnaire

Question 11.04 :	Yes	No	N/A
Do you make documentation on your sources and methods publicly available?	74.4	18.2	7.4

Metadata

3. Result from world wide consultations

Question 15: Do you agree that IMTS, Revision 3, should provide additional recommendations on the compilation and dissemination of metadata as an integral part of the regular work programme of trade statistics compilers?	**Yes**	**No**	**No Op.**	**N/A**
Total of 110 replies	89	5	4	2
Developed (34)	88	9	3	0
Developing and transitional (76)	89	4	4	3

4. UNSD observation of wwc result

There is universal support (89%, with no-answers 5%) for this proposal. Many comments suggest providing a list of items or a template, preferable in the Compilers Manual. USA warns that agreement on details will be difficult to achieve and a few countries say there is not much need for such additional recommendations. Also reference is made to existing reporting standards such as the IMF's GDDS or the reporting requirements within the EU.

Metadata

5. Selected comments (from total of 39 comments)

- "The current recommendation, to our mind, already states the principles and preferred practices quite clearly."

- When comparing different countries' trade data for economic monitoring and trade policy purposes, it is useful to have an overview of the methodologies used to inform analysis."

- "The recommendation would also need to include timeframes for updates and release of information."

- "This is a quality issue"

Experiences from UN Comtrade

Current situation

- Collection of crucial metadata is often very difficult or impossible to obtain – also the majority of countries do not include metadata in their data submissions to international organizations
- The same metadata information from difference sources sometimes contradicts each other.
  - For example, trade system from NCDP differs with trade system mentioned in the website. Any attempt to clarify is sometimes in vain.

Available metadata in UN Comtrade

- Explanatory notes which consist of
  - Trade system, valuation, partner country attribution, original currency, conversion factor and original reported commodity classification.
- Publication notes which consist of
  - Date of receive, data provider, date of release, information regarding data revision or any important processing notes

Experiences from IMF GDDS/SDDS

What is GDDS/SDDS

- Both are IMF's initiatives on data dissemination standards
- GDDS: General Data Dissemination System; SDDS: Special Data Dissemination System
  - SDDS has more stringent requirements (shorter periodicity and strict timeliness)

Relation with Data Quality Assessment Framework (DQAF)

- DQAF grew out of GDDS/SDDS
- The SDDS and GDDS metadata components are alternative aggregations of the DQAF's *indicators*.
  - Oriented to dissemination practices
- There is work in progress of integrating SDDS/GDDS metadata to DQAF structures

Experiences from IMF GDDS/SDDS

GDDS/SDDS Components

- Data characteristics
  - Coverage, Periodicity, and Timeliness
- Quality
  - Documentation of Methodology
  - Data to Support Cross-checks and Assurance of Reasonableness
- Integrity
  - Terms and Conditions Under Which Data are Produced and Disseminated
  - Identification of Internal Government Access to Data Before Release
  - Provision of Information about Revisions and Advance Notice of Major Changes in Methodology
- Access
  - Simultaneous Release to the Public
  - Advance Release Calendars

Possible issues for discussion

Data Quality

- Due to universal support of adding recommendation of data quality/metadata, it is planned to add chapter "Data Quality & Metadata" in the next revision.
  - Should it contains example of quality framework or dimensions and indicators for data quality?
  - Should it deals with practical steps / measures / best practices?

Metadata

- What are the possible main elements of metadata that should be compiled and disseminated as necessary part of regular work programme of trade statistics compilers?
- How to reconcile conflicting metadata information from different sources?
- Comments on the use of NCDP as basis of metadata on IMTS.

Thank you for your attention