

Project 06/07K: Strengthening the Development of International Merchandise Trade Statistics (IMTS) and the compilation of e-commerce in Member Countries of the Economic and Social Commission for Western Asia (ESCWA)

Workshop on the compilation of IMTS and the application of methodological concepts

28 October – 2 November 2007, Amman, Jordan

Data Validation :

Use of technology

Overview

by Markie Muryawan

Statistician

Current Data Validation Practice

- It is observed that validation by experience (which depends on person) is practiced by some countries
- It is true that experienced customs officers or statisticians can be very effective in detecting errors, however:
 - They have limited time (only work days)
 - They can be overwhelmed by high work load, which lead to decrease in their effectiveness
 - They may not be available or replaced by less experienced person

Validation Tools

- Validation tools are not intended to replace resource person

but

- Just like other tools, it served to increase effectiveness and productivity of resource person
- Major errors, to be verified further, can be identified faster and more accurately
- In addition, automatic data correction (for basic error) can help reducing workload of resource person

IT as validation tools

- IT improvement in past years has enabled the use of technology for validation of “very large” data can be done in relatively inexpensive system such as PC
- It is worth to mention that the main problem in the implementation is not availability of tools

but:

- How to reform/change established working culture?
- What is the right software for data validation?
- Has the software implemented best practices and national validation procedures?
- Do staffs have enough training?

Code Validation & Mandatory Value

- Current major database products (MS Access, Oracle, MS SQL, Sybase, etc) are able to:
 - Enforce referential integrity (it means that codes must be on reference tables otherwise it can't be stored)
 - For an example, if there is no "Europe" in reference table "Partner Country", any data that refers to Europe will be rejected
 - Set specific data field to be mandatory
- The first step of validation is to filter all non-standard codes and missing values, then to create rejection reports for further clarification

Unit Value Validation

- Unit value filtering can catch registration errors related to value and quantity (too high/low)
- As prerequisite, it is important to have reference unit value ranges as basis for filtering (standard or historical unit values)
- Reference unit values can be stored in database
 - Unit values can also important for other purposes, such as development of trade indices
- Report should be created to show all rejected unit values
 - It is also interesting to show % of rejected records from time to time

Partner Country Validation

- In addition to code validation, partner country can also be verified based on combination of commodity/trade flow/partner (such as no export of banana from Sweden)
- These kind of validation are built by experience, so by capturing the knowledge in the system, the chance of losing the knowledge is minimized
- Experience can differ from country to country

Example of data validation tools

Eurotrace 2003
Version Windows

EUROTRACE

The logo for EUROTRACE features the word in a bold, blue, sans-serif font. The letter 'O' is a solid yellow circle. Several yellow five-pointed stars are scattered around the letters, with a cluster of seven stars positioned to the left of the 'EURO' part of the word.

External Trade Statistics Software

The Next Generation Upgrade

Version 2003

Data Processing Software for External Trade Statistics

Three separate applications that work together.



Eurotrace DBMS



Eurotrace Data Editor

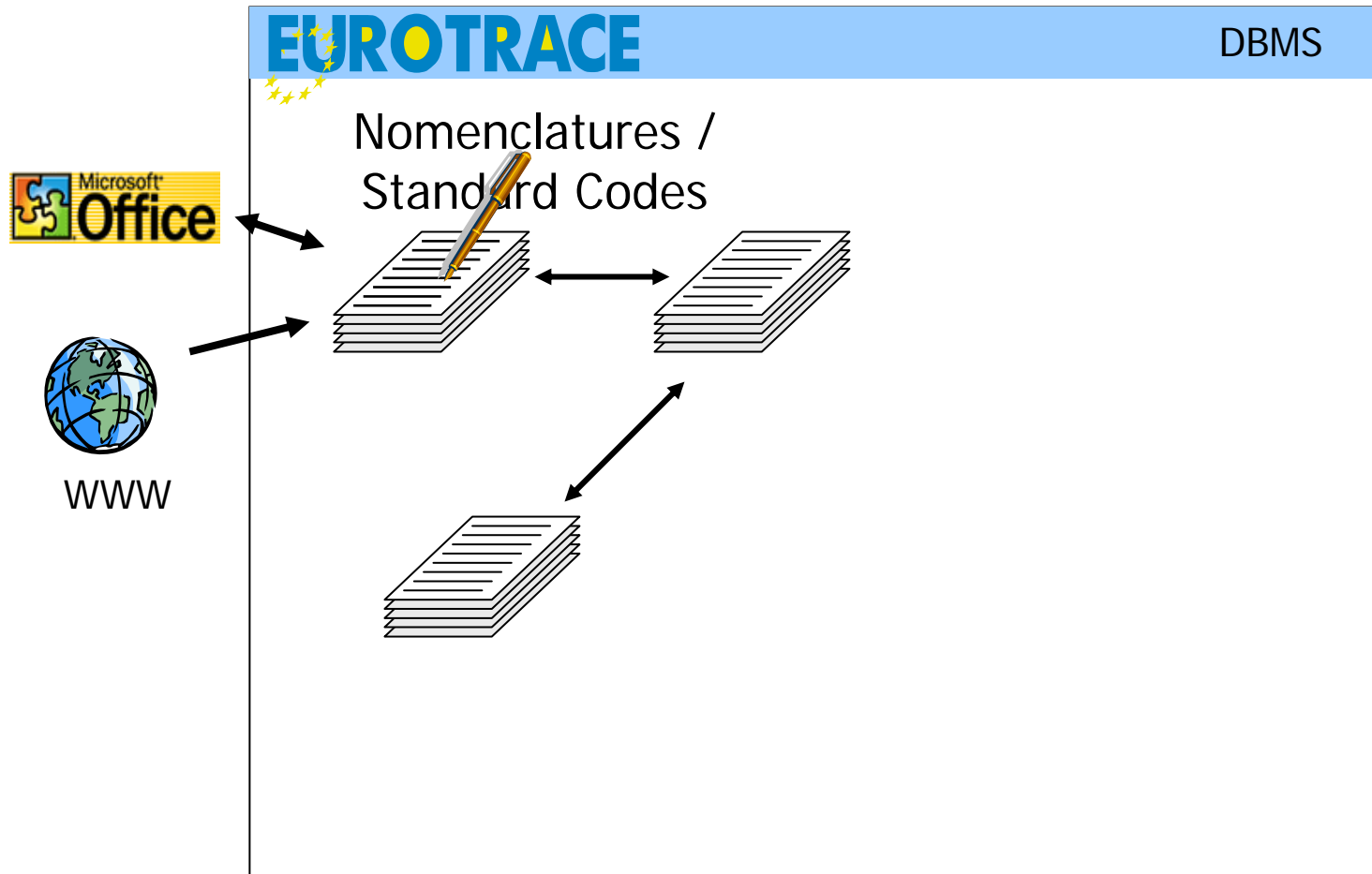


Comext Standalone Data Browser

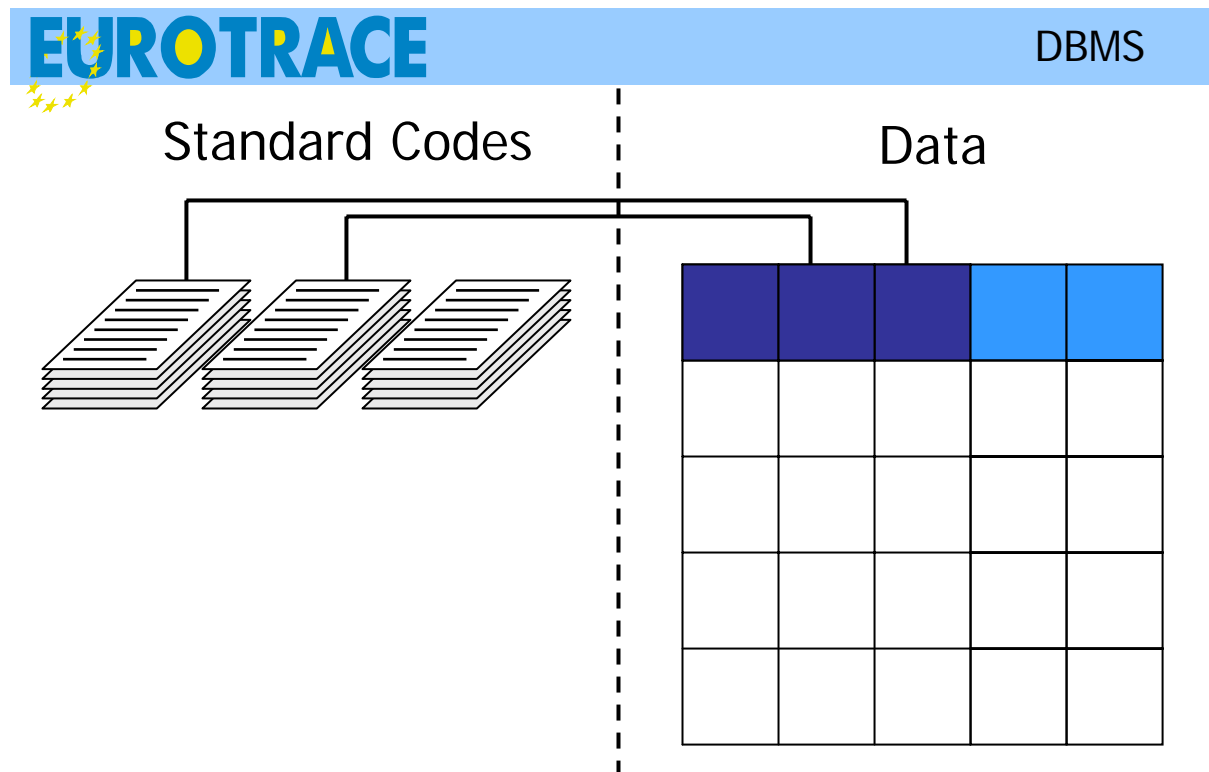
Available **Free** to participating countries.

International Trade Statistics Section, October 2007

Nomenclature Management

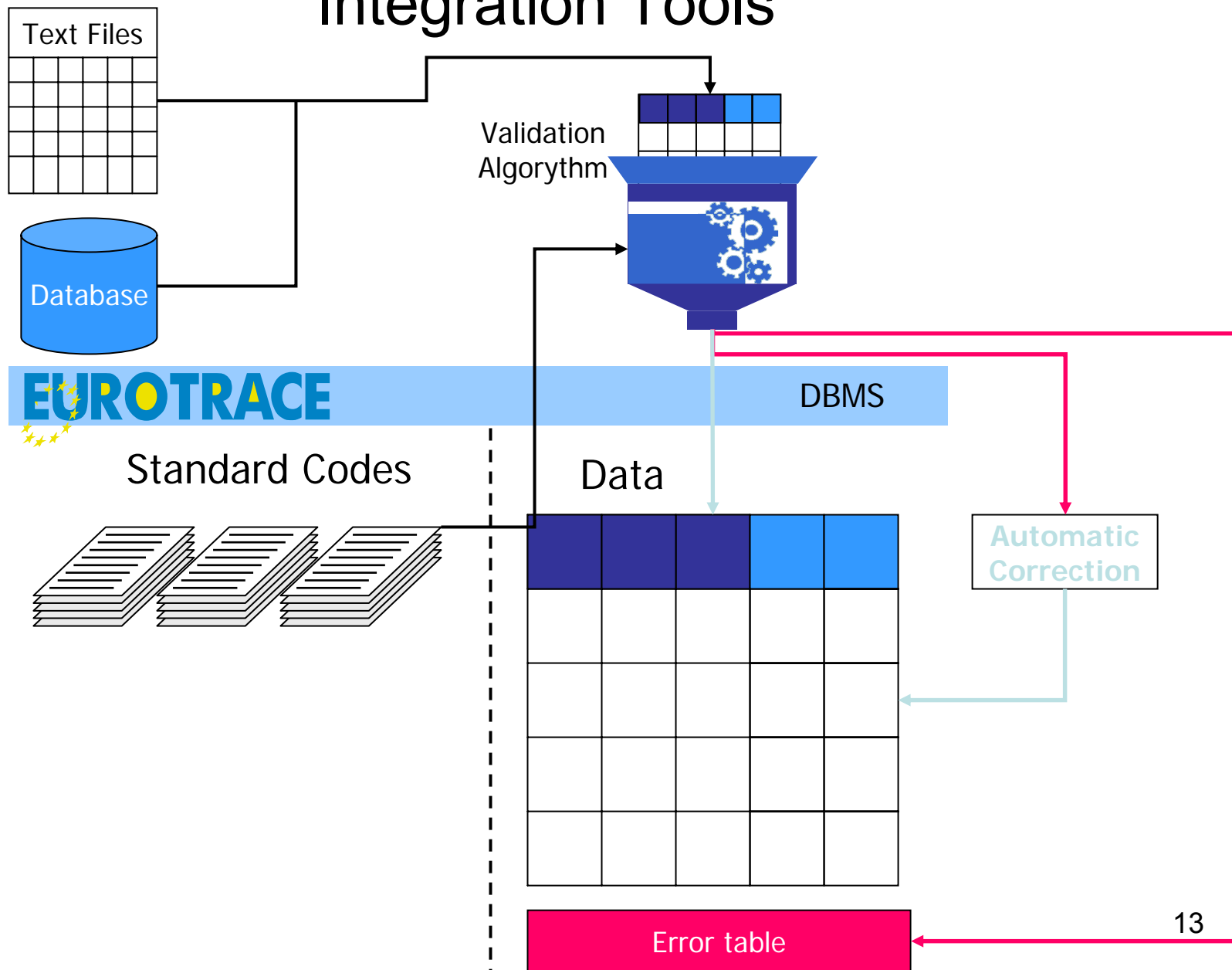


Flexible Storage Structure

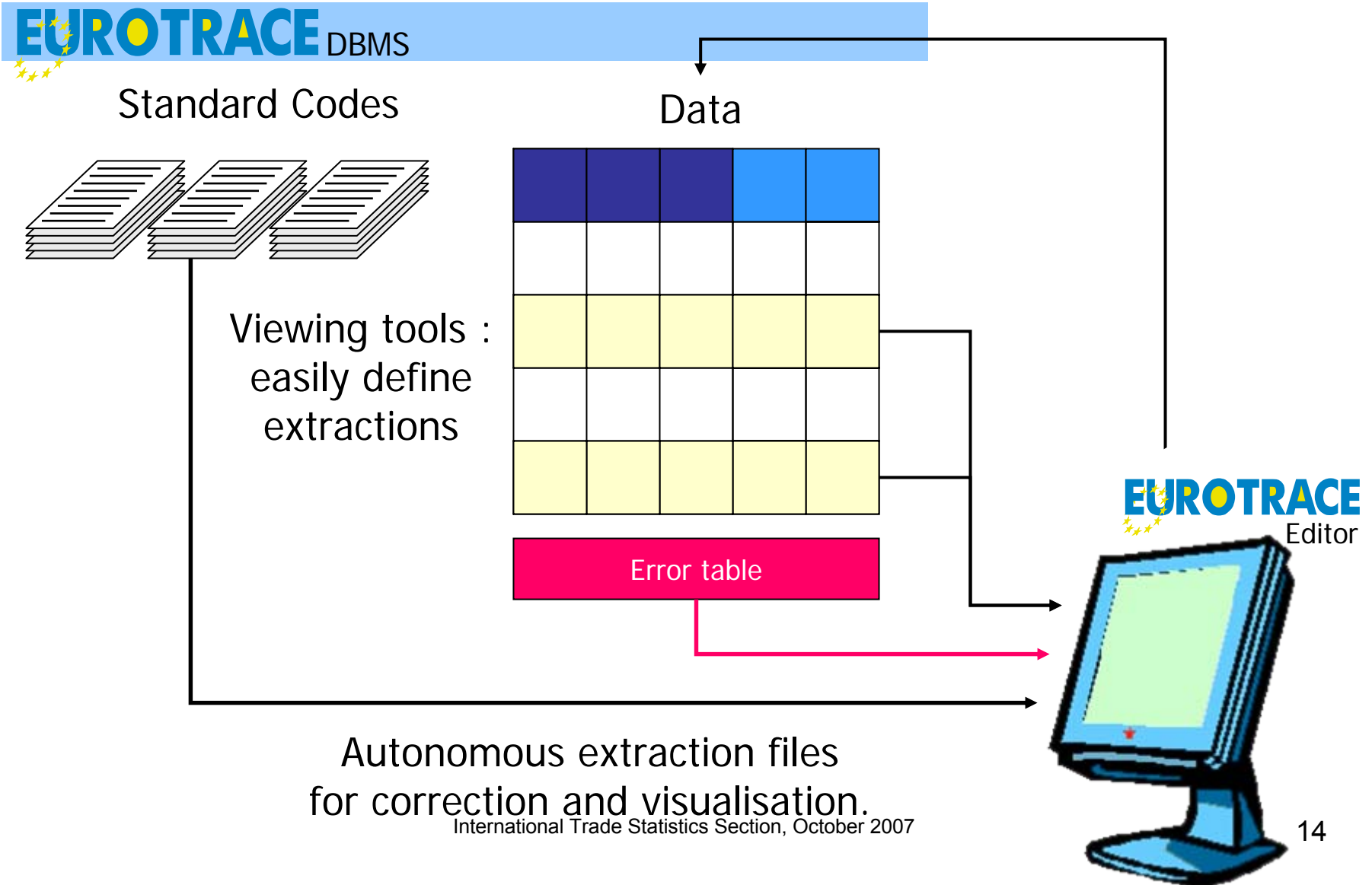


Unlimited number of dimensions, ex: custom procedures/ports/partner/etc

Integration Tools



Transfer Tools

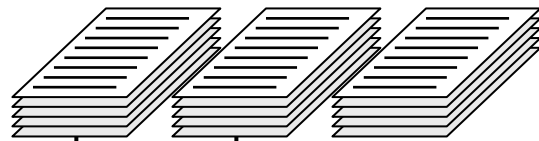


Aggregation of the data

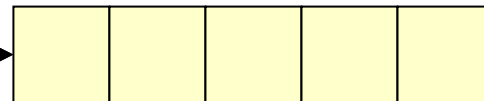
EUROTRACE DBMS

Standard Codes

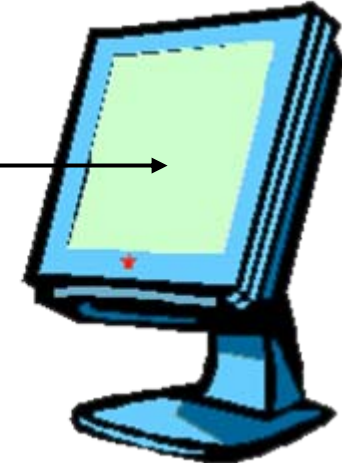
Data



Aggregates



COMEXT
Browser



وشكراً