

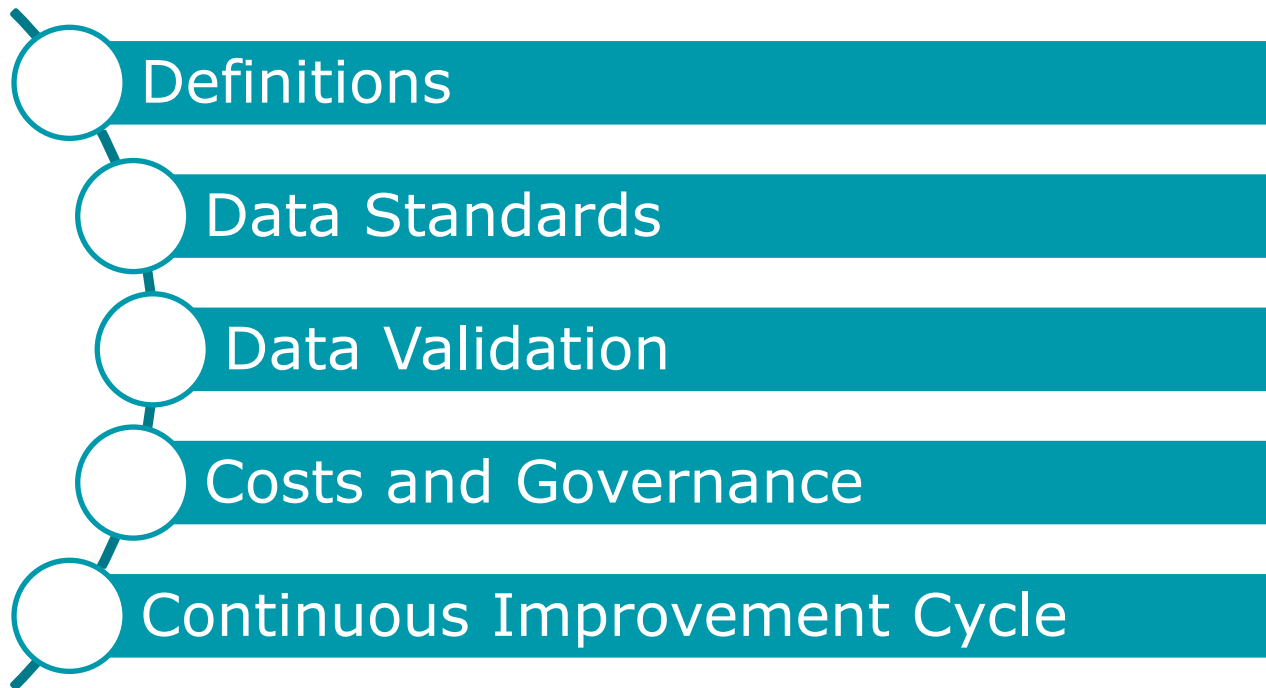


Data Quality: An Introduction

Stefano Cappè

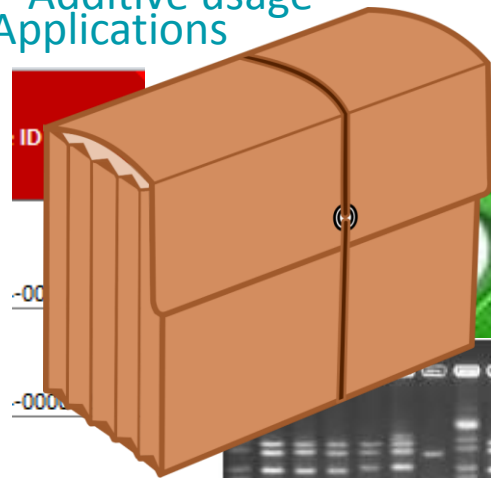
EFSA Stakeholder Forum
Parma, 30 – 31 May 2017

DATA QUALITY

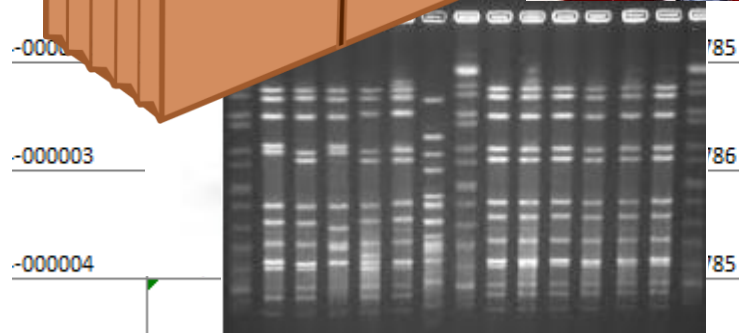


DEFINITIONS: DATA

Additive usage
Applications

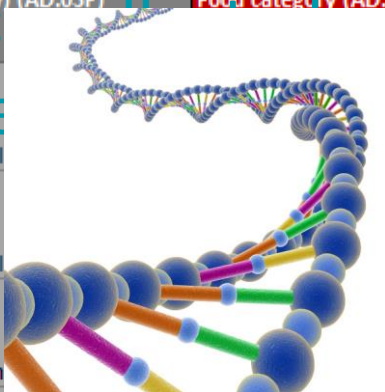


Images



Molecular Typing

Oxford dictionaries

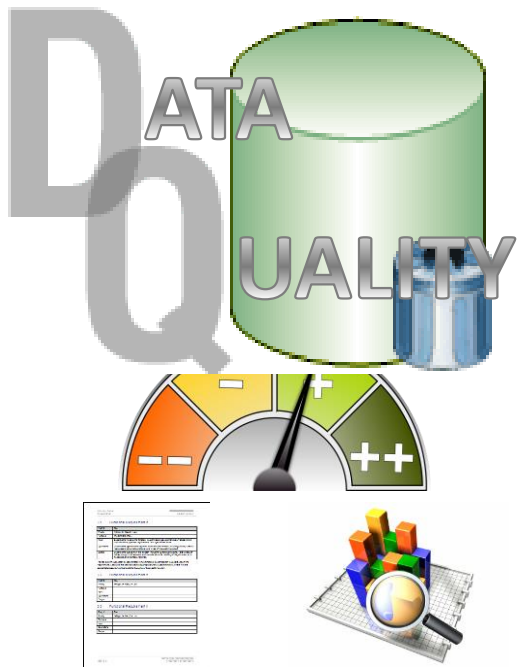


Whole Genome Sequencing

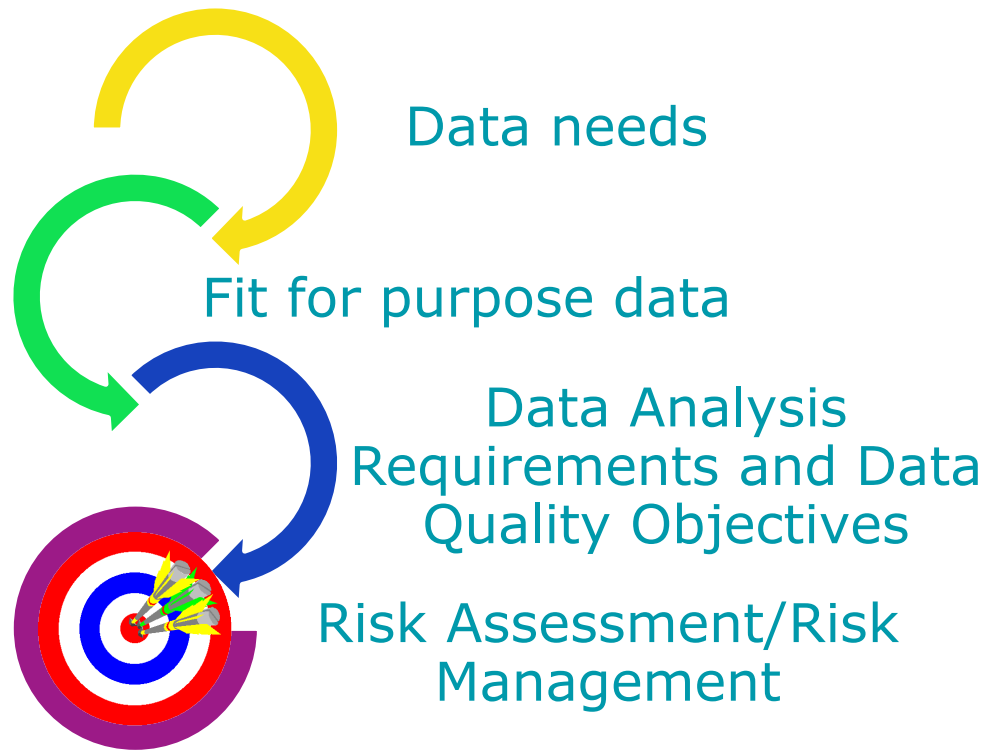
Documents

Food category (AD.04)
for special medicinal purposes (excluded category 13.1.5)
for special medicinal purposes (excluded category 13.1.5)
for special medicinal purposes (excluded category 13.1.5)
13.2 Dietary foods for special medicinal purposes defined in Directive 1999/21/EC (excluded products from food category 13.1.5)
13.2 Dietary foods for special medicinal purposes defined in Directive 1999/21/EC (excluded products from food category 13.1.5)
13.2 Dietary foods for special medicinal purposes defined in Directive 1999/21/EC (excluded products from food category 13.1.5)

DEFINITIONS: DATA QUALITY



Distance between data needed and
data available/obtained



DATA STANDARDS: SOME EXAMPLES...

Contaminant
Occurrence

Additive
Occurrence

(revision 2)

Veterinary
Medicinal
Products

Pesticide
Residues

hierarchies were added and the relationship between the terms and the most important facets was streamlined. This technical report, mainly aimed to data providers to EFSA, describes the revision of

DATA STANDARDS: SOME EXAMPLES...

Globally Harmonised Submission and Transport Standard

ENV GHSTS

[Home](#)[Introduction](#)[GHSTS downloads](#)[Documentation & Training](#)[Frequently asked questions](#)[News](#)[Roles & responsibilities](#)[Archived version](#)

ENV GHSTS

About the Globally Harmonised Submission and Transport Standard

The Global Harmonised Submission and Transport Standard (GHSTS) is used to assemble electronic data files in PDF, and XML data files in business needs. Once the data is transferred from a business to a regulatory authority, the GHSTS can potentially

Application dossiers

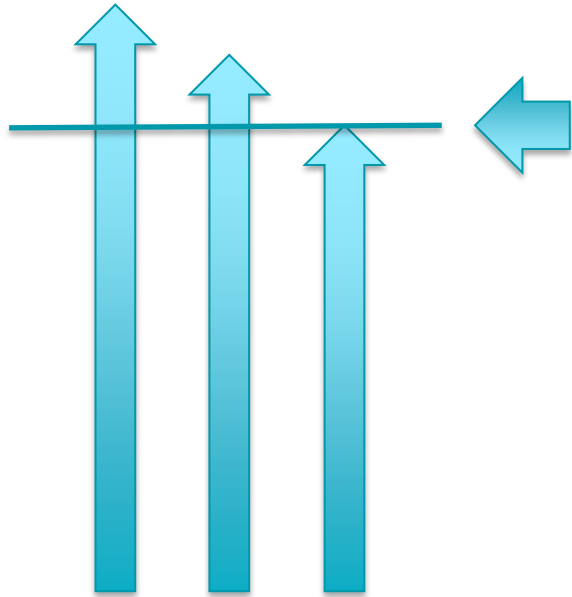
of technical specifications in Microsoft Word and Excel, Adobe PDF, and GHSTS according to the GHSTS standard. The GHSTS is transferred from a business to a regulatory authority in a regulatory process. The GHSTS can potentially be used as well.

Latest news

[The GHSTS version 01 is published May 2014.](#)

DATA VALIDATION: SETTING MINIMUM QUALITY

Different data uses requires different quality



Data validation to standardise minimum quality levels

Same data standard but validation rules are data specific (e.g. for contaminants, for additives)

COST AND GOVERNANCE

Standard Sample Description

Sample identifier
Country of sampling
Food description (coded)
Food description as text
Food treatment
Analytical method
Analytical result value
Limit of quantification

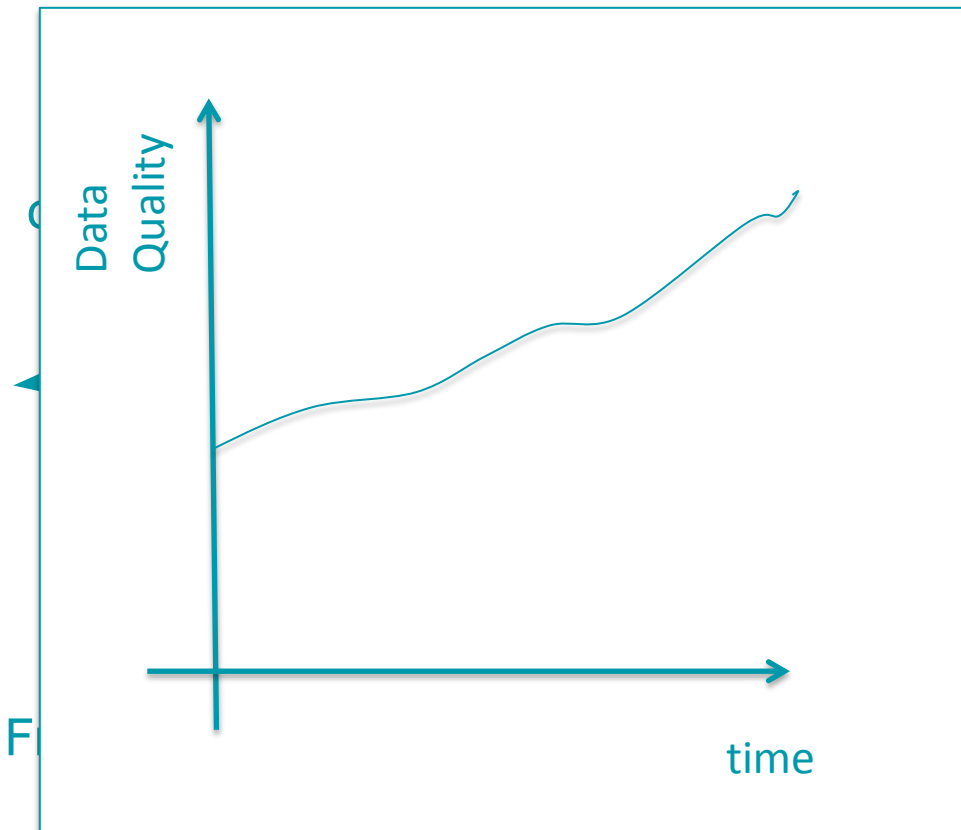
Each data element has a cost



Data Suppliers Costs vs Data Users Needs

E.g. National Competent Authorities discuss validation rules in Scientific Networks to improve data quality

CONTINUOUS IMPROVEMENT CYCLE



Economical Methods



Economical Data management



More data and more detailed data (Big Data)



More societal expectations



More data quality needs

Data Quality embedded at source:
Data generation – Data entry

CONCLUSIONS

Data Standards are the pillars of Data Quality

Data Quality improvement and data validation discussed with EFSA Data Providers

Data Quality expectations increases with time

Data Quality must be embedded in EFSA Data Providers processes, from data generation/data entry

THANK YOU



Questions?