40th Focal Point meeting 4-5 September 2019

Future data collection improvements & involvement of Member States

Claudia Heppner

Evidence management unit

Trusted science for safe food

efsa
European Food Safety Authority

Focal point (FP) grant agreement 2019-2022: DATA matters

5. Support on data related matters	
Specific activities	***Priority***
1. FP maintains an up-to-date list of national data providers across data domains (areas) relevant to EFSA's remit[17].	*low*
2. FP supports EFSA in facilitating two-way communication with national data providers on a need basis[18].	*low*
3. FP supports the promotion of best practices for sharing public sector information. To this end, and if needed, the FP shall identify a national contact point to support the implementation of this activity.	*high*

Background:

- Data is not collected for multi-purposes, and not only for EFSA; hence data is not always in the EFSA format
- Fragmentation of the services for the data entry (from sampling to analysis results) up to EFSA submission
- Data quality is important. What is data quality?

  Data quality is the extent to which data are **fit for their intended use**[1)]
- In 2017, a pilot framework partnership grant agreement on data quality (50% co-funding from EFSA) was launched with five MS (Cyprus, Germany, France, Denmark, Slovak Republic) for the following domains: contaminants, veterinary medicinal product residues, pesticide residues and zoonoses monitoring data.
- Objectives - pilot:
  1. Improve data collection co-ordination at national level;
  2. Monitor and improve data quality by define and measure data quality objectives (for incoming data to EFSA); and implement agreed technical enhancements to national reporting systems

1) In line with the general definition of quality as set in the standard ISO 9000: "degree to which a set of inherent characteristics of an object fulfils requirements

Background – cont.:

- Quantification of data quality – dimensions

Validity

- Are data elements consistent to their format, type and range?
- Are constraints respected?

Uniqueness

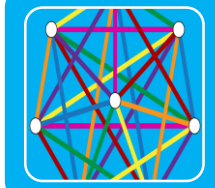
- Are the records present only once in the database?
- Are database unique identifier available?

Timeliness

- Are data available when needed?
- Are data up to date for their uses?

Accuracy

- Are data elements representing correctly the real world from which are extracted?
- Are the data plausible?



Completeness

- Is information reported in the data elements comprehensive?
- Are valuable data elements missing?

Consistency

- Are different data elements providing non-conflicting details for a specific piece of information?

Background:

Example of quality indicators from contaminants data domain

Data quality objective	Data quality performance indicator	Data quality dimension
Timely transmission of data	Proportion of data records in "SUBMITTED" status by data collection deadline	Timeliness
No duplicate records submitted	Proportion of data records not duplicated	Uniqueness
Correct identification (coding) of the food/feed matrix analysed	Proportion of records containing consistent information in SSD data fields identifying the food/feed matric analysed	Accuracy
Correct identification (coding) the chemical substance analysed	Proportion of records containing consistent information in SSD data fields identifying the chemical substance.	Accuracy

Overall score is calculated by dimension using full dataset or a random sample

Background – cont.:

- New 178 Transparency Regulation (revision of the General Food Law (Regulation 178/2002) lays down provisions on quality and reliability of studies, standards of data formats, and governance and sustainability (cooperation with MS)

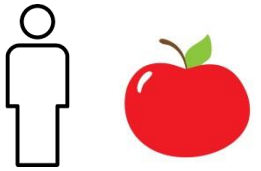
Questions to data providers via FP

- **Which are the major issues impacting on data quality when submitting data to EFSA?**

- **If funding would be available, what type of activities would data providers/MS like to have funded to resolve the issues?**

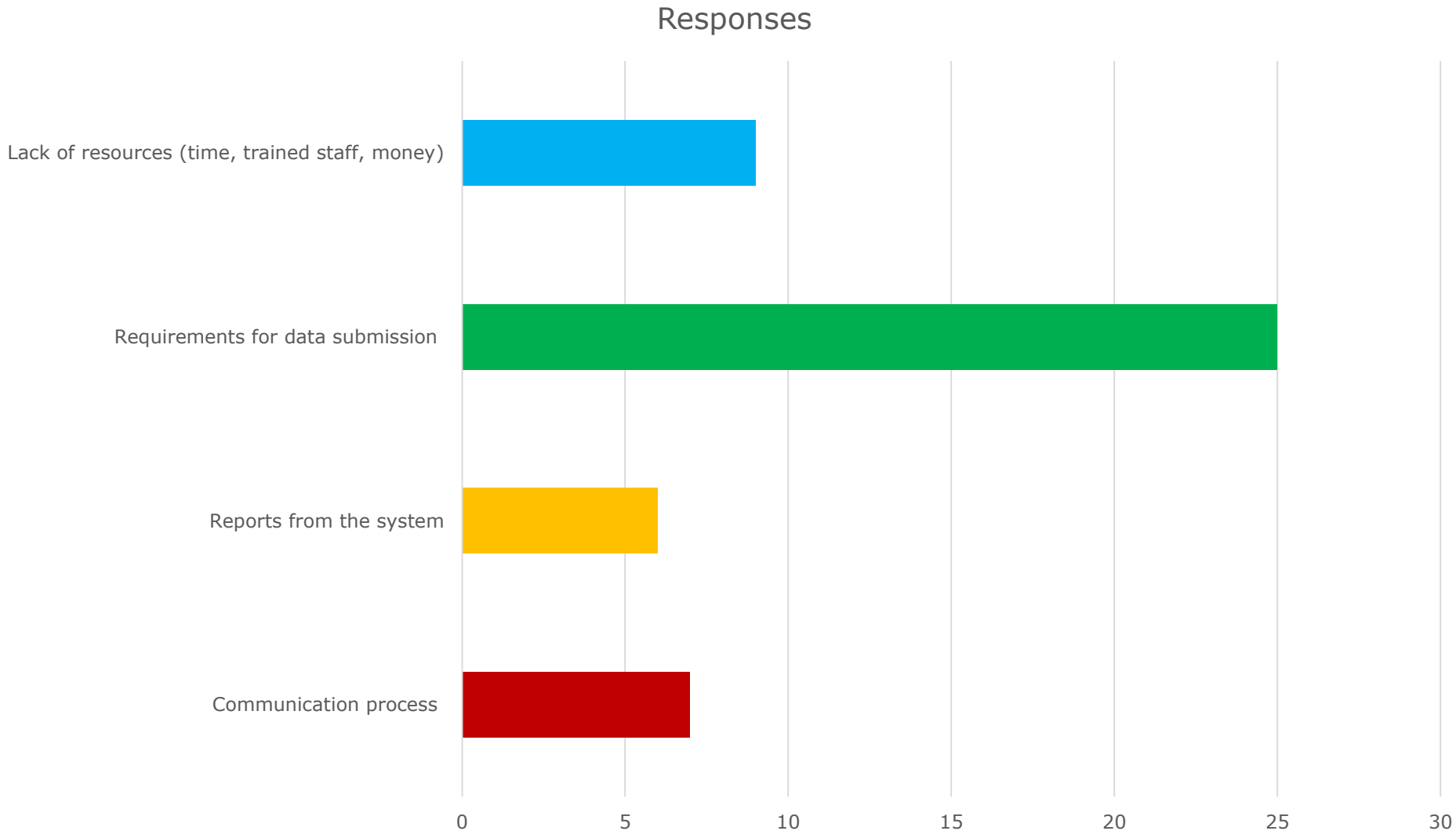
Occurence data

Consumption data



Responses from 27 FP

Major issues: summary



Major issues: requirements for data submission

Issue	Description
Data submission	frequent changes; too short time for implementation (2y needed); many information, mandatory fields; not always in the GD or no specific catalogues)
	Reporting is complex (complex data models and guidance documents)
	Format and data conversion (mapping of FoodEx2 (differences in FoodEx1 and FoodEx2; not coded on receipt)
	Maximum levels for all domains are not available in one place
	Often analysis are not accredited
	Information requested and timing (too much not available at national level, difficulty to categorize a product arriving to the laboratory according to EFSA CAT MATRIX)
	Not real-time data entry
	Manual recording (no LIMS in place, data not fully standardised in database)
	IT issues (DP can not download information; RAM insufficient, EFSA BR are not included; maintenance and updating of system funding)

Major issues: resources

Issue	Description
Lack of resources (time, staff, money)	Lack of time : - to prepare reporting; - For quality check;- to adapt national system to new rules,
	Time shortage: Although there is a LIMS system in place, it is not designed for preparing aggregated data. Therefore, statistical analysis and preparation of data is still carried out, to a large extent, manually, resulting in a highly time-consuming procedure
	No standing committee of staff dedicated to the task of compiling & processing the data, for onward transmission to EFSA.
	Understaffed services
	Lack of IT support/funding for data submission/whole programme
	Correction of data and feedback – time consuming

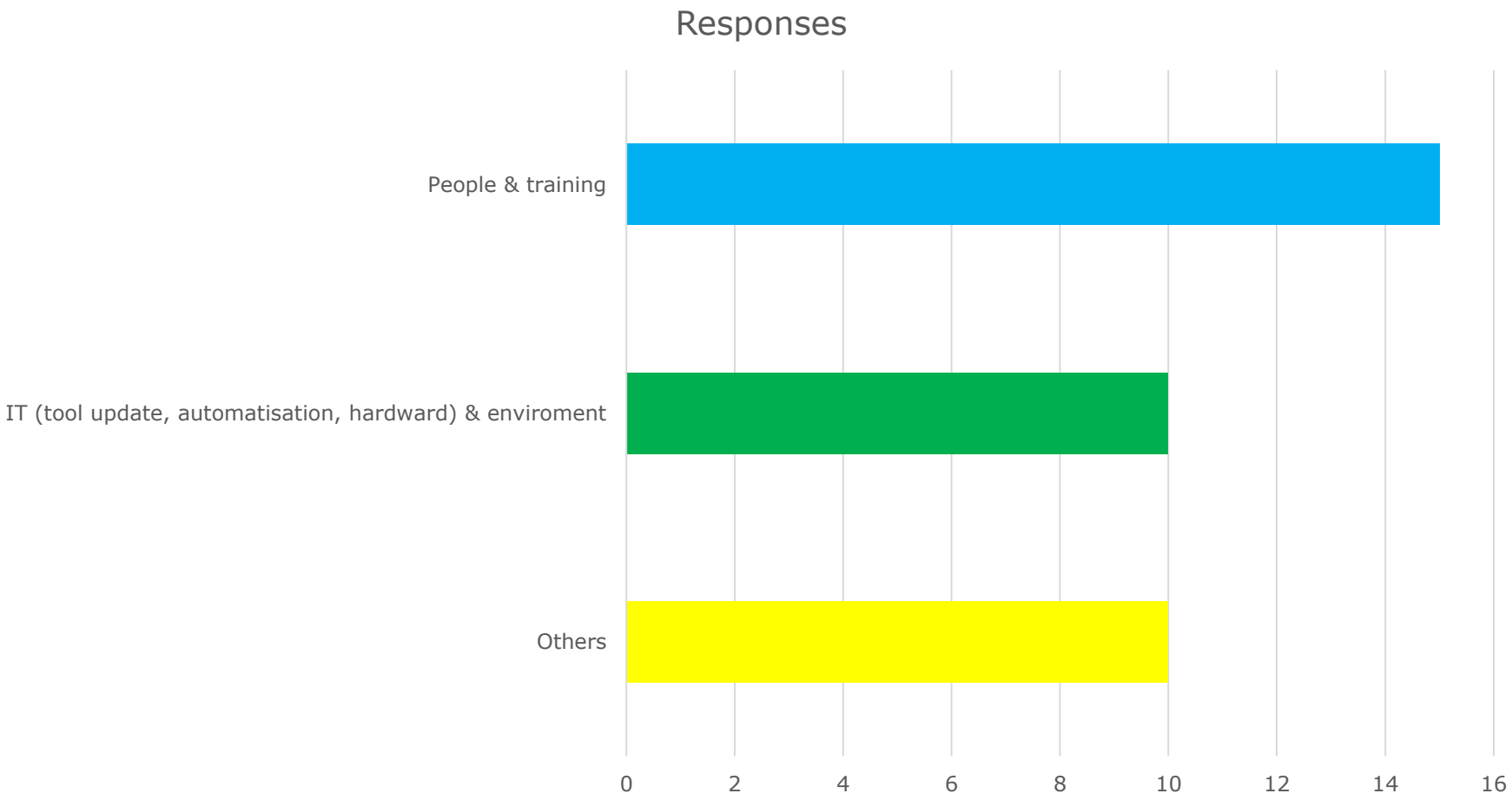
Major issues: reports from the system

Issue	Description
Reports (acknowledgement, validation)	complex not easy to understand: understanding the encrypted acknowledgement (ACK) messages; when opening them in a XML format a long list of mixed codes is received and makes it complicated to find a mistake
	not user-friendly;
	difficulties (zoonoses RO) to follow the data in the final (summary) tables and reports prepared by EFSA, available in the data warehouse, and to trace back to the data sources because of different format and visualization of the aggregated data)
	Unclear feed-back from automatic control system (what need to be corrected) – SSD2
	The data to be corrected received from automatic control system are sent only in .xml format that is difficult to view (maybe it will be good to send also in other format (word, excel, pdf);

Major issues: communication

Issue	Description
Communication process	Lack of awareness of EFSA DQ at laboratory level, communication (of changes) and implementation at input level
	When changes related to the SSD2 system occur a request to send a short letter, where the essential points are stated clearly to all data providers, with an indication in the subject field - IMPORTANT - would be appreciated
	The labs need to be aware of EFSA's stringent data quality requirements at the input level
	For MS remains still difficult to understand, which EFSA-platform is used for each type of information. Some information from our colleagues in sister institutions in Germany were sent few months ago
	Better solution for the text forms (to many subtitles). Word Documents are no longer adequate

Proposed activities: summary



Proposed activities to solve issues: People

Issue	Description
People & training	Support for the use of the "Harmonised Reporting Guideline
	Coordination on technical and scientific questions regarding data collections
	More personal support of EFSA for MS data providers to ensure good reporting and feasibility of new rules
	There should be constant human resources to make those changes to our systems
	National reporting expert/officer to manage and coordinate country-wide reporting. Task could comprise: • participate in EFSA´s networks; coordinate reporting; act as reporting expert and support data providers related to data; develop of reporting/mapping tools together with EFSA; submit data to EFSA; train national DB
	there is a need for dedicated staff and/or sample based reporting combined with upgrading of the LIMS system
	training is required for DPs on how to report data using different EFSA tools and databases either staff on-site (locally), or at EFSA facilities
	Training on sample description of the staff involved in the different steps of the data collection and transmission

Proposed activities to solve issues: IT

Issue	Description
IT (tool update, automatisation, hardward) & enviroment	IT Systems upgrade which are used in data collection and processing
	Modification, or upgrading, of database software
	Tools to implement national data's in EFSA format and updates of the tools (advanced electronic tools)
	High performance computer hardware to rapidly process the bulky data files.
	tool so that Foodex coding can be applied at sampling point, this would greatly improve accuracy and quality of data; IT tool for automatic assignment of Fodex2 codes
	Funding for testing phases for new or changed data collections
	Improvement of automated data reprocessing processes (computer programming
	Funding of projects for barcoding of food samples at nation level

Proposed activities to solve issues: Others

Issue	Description
Others	Funding for testing phases for new or changed data collections
	More feed-back about the use of the data's by EFSA
	to coordinate with the ministries to improve the quality of the data at source (the data we receive from the ministries) as the data are not clean
	table with MLs contains Foodex2 name, the parameter to be analysed and the name of the Directive from which the ML values are taken
	Funding to support reporting data on additives and other domains that are not mandatory yet
	Food consumption: updating the recipes at national level and harmonizing food composition data at EU level.
	EFSA to consider collecting data on food contact material.
	Designing our internal systems to be easier to manage and align with EFSA data standards i.e. Database, SSD2 catalogues
	Reopening of the framework partnership agreements with Member States on data quality: for optimising the processes of data collection and submitting data to EFSA, Lean process for EFSA data collection on a national level

Proposed next steps

- DATA unit will align feedback received with data networks feedback/discussions;
- FP to develop a national action plan related to improving data quality for timeliness and in alignment with AF TF data collection;
- FP to explore opportunities for training and knowledge sharing related to data transmission;
- EFSA is currently exploring how to redesign our data processes under the new Transparence Regulation


DATA quality

claudia.heppner@efsa.europa.eu

Acknowledgement:
Mary Gilsenan, Eileen O´Dea, Jane Richardson, Stefano Cappe
Evidence Management unit

Thank you for your attention