GERMAN FEDERAL INSTITUTE FOR RISK ASSESSMENT

BfR
Bundesinstitut für Risikobewertung

Improvement of the OECD Transport Concept

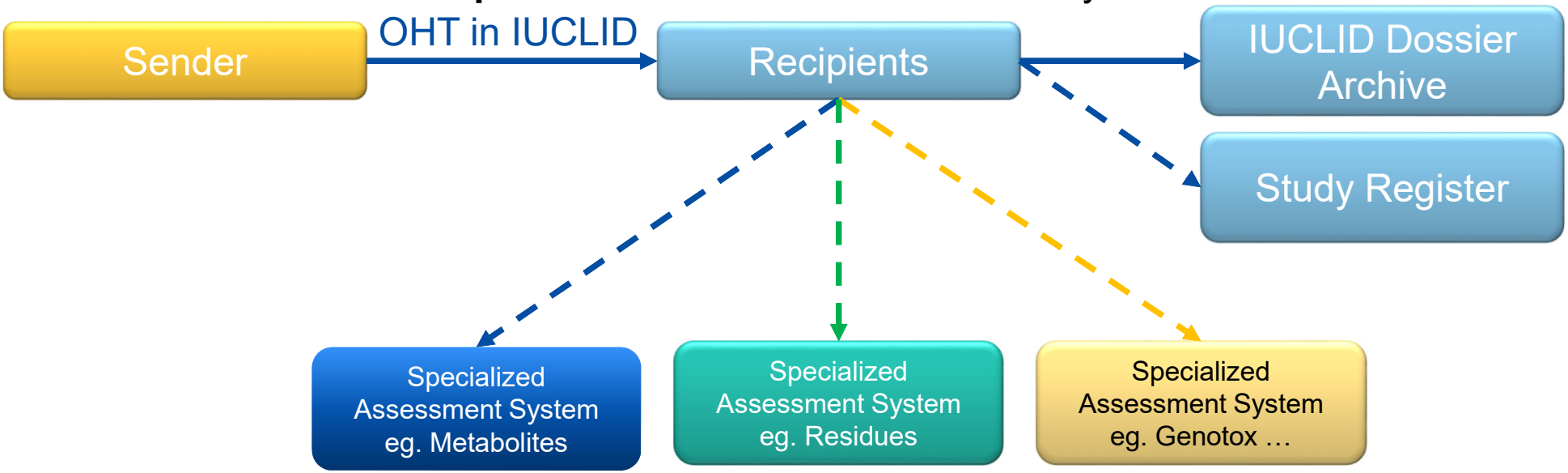
Stephan Worseck

Who are potential recipients of aggregated raw data ?

The OECD implemented OHT schema for information transport on a **semantic level**

The receiver and sender can agree to use IUCLID as a **transport system** for OHT's

The receiver can decide to **process the received data** in subsystems



Transport concepts for aggregated raw data of **metabolism studies**

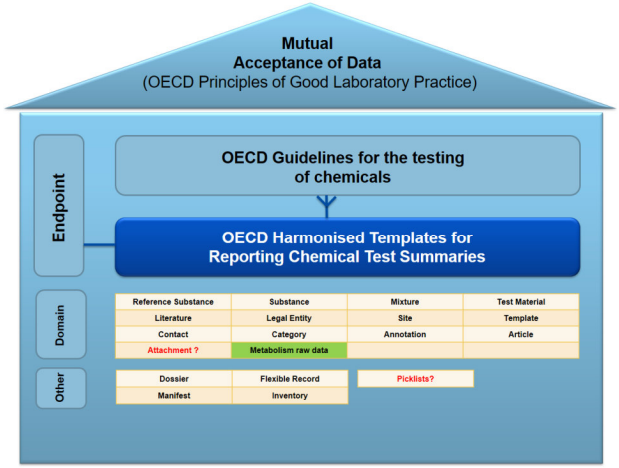
The solutions should guarantee the required information flow in appropriate granularity.

Only technical solutions that are **compatible with the IUCLID submission approach** are considered for the transportation step.

The following aspects should be included in the decisions:

- Who is the addressee for the information flow of the required **aggregated raw data** – the human or a machine?
- The effort to initialize this information flow with all the needed IT-Systems and data interface,
- the effort required to maintain the systems and
- last but not least: the solution should be usable world wide, **also without IUCLID.**

The option “Create an OECD Domain Type” for all metabolism studies , which is **not** recommended by BfR

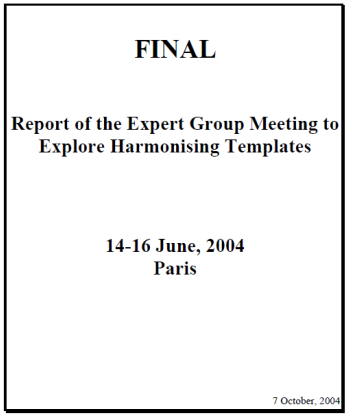


All technical problems can be solved with reasonable effort. The implementation of an “OECD Domain Type” from metabolism studies is possible.

This domain type should then be generic enough to be included in all “OECD Harmonised Templates” for metabolism experiments, similar to the domain type “Literature” which can be included as an input frame for each “Data Source” section.

But there is a high risk that the OECD will destroy it’s own transport concept!
The Explanation follows in the next slides.

The old compromise between structured metadata and free text fields (I)



FINAL

Report of the Expert Group Meeting to Explore Harmonising Templates

14-16 June, 2004
Paris

7 October, 2004

FIELD TYPES

FIXED TEXT			FREE TEXT (w/ subfields)			FREE TEXT (w/ guidance)		
Field No.	**Field Name**	**Subfield**	**Field No.**	**Field Name**	***Subfield***	**Field No.**	**Field Name**	**Subfield**
			. . .			. . .		
Animal			Animal			Animal		
5	Species	Picklist ▼	5	Subheading: • Species *Enter Text* • Strain *Enter Text* • Sex *Enter Text*		5	Guidance: (Species, Strain, Sex) *Enter Text*	
6	Strain	Picklist ▼						
7	Sex	Picklist ▼						

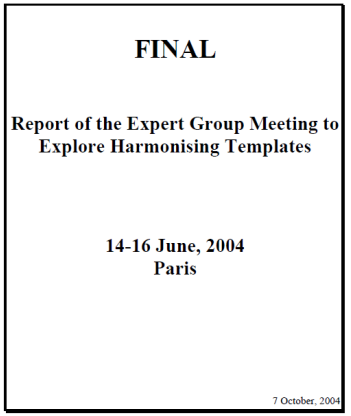
High Searchability — **Low Searchability**

Easy to manipulate or further process data — **More difficult to manipulate or further process data**

Difficult to migrate old data into — **Easy to migrate old data into**

FINAL Report of the Expert Group Meeting to Explore Harmonising Templates 14-16 June, 2004 Paris 7 October, 2004

The old compromise between structured metadata and free text fields (II)


FINAL

Report of the Expert Group Meeting to Explore Harmonising Templates

14-16 June, 2004
Paris

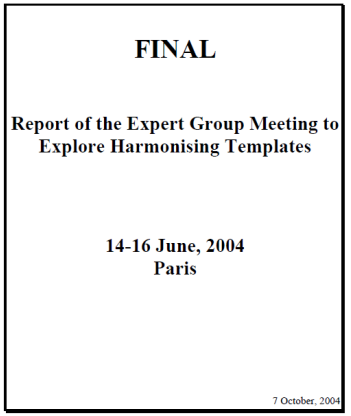
7 October, 2004

The type of field included (fixed-field versus free text), should:

i. be based on the needs of the reviewer and not the electronic technology requirements;
ii. consider how often that field will be searched and by whom (i.e., searching is easier with fixed-fields than free-text fields);
iii. consider the need for future manipulation of both text and numeric data in specific fields, e.g., extracting text blocks and/or numeric data into evaluation reports, performing statistical analyses, data mining, or other mathematical operations. For these tasks, fixed fields generally provide a greater ability than free-text fields;
iv. Consider whether (and the degree to which) old, unstructured free-text data will be migrated into a field (migration to free-text fields is easier than to fixed-fields)

It seems that the first condition, "**be based on the needs of the reviewer and not the electronic technology requirements**" has been increasingly forgotten over the years, and more and more subject-based metadata has been incorporated into OHTs.

The old compromise between structured metadata and free text fields (III)



FINAL

Report of the Expert Group Meeting to Explore Harmonising Templates

14-16 June, 2004
Paris

7 October, 2004

The type of field included (fixed-field versus free text), should:

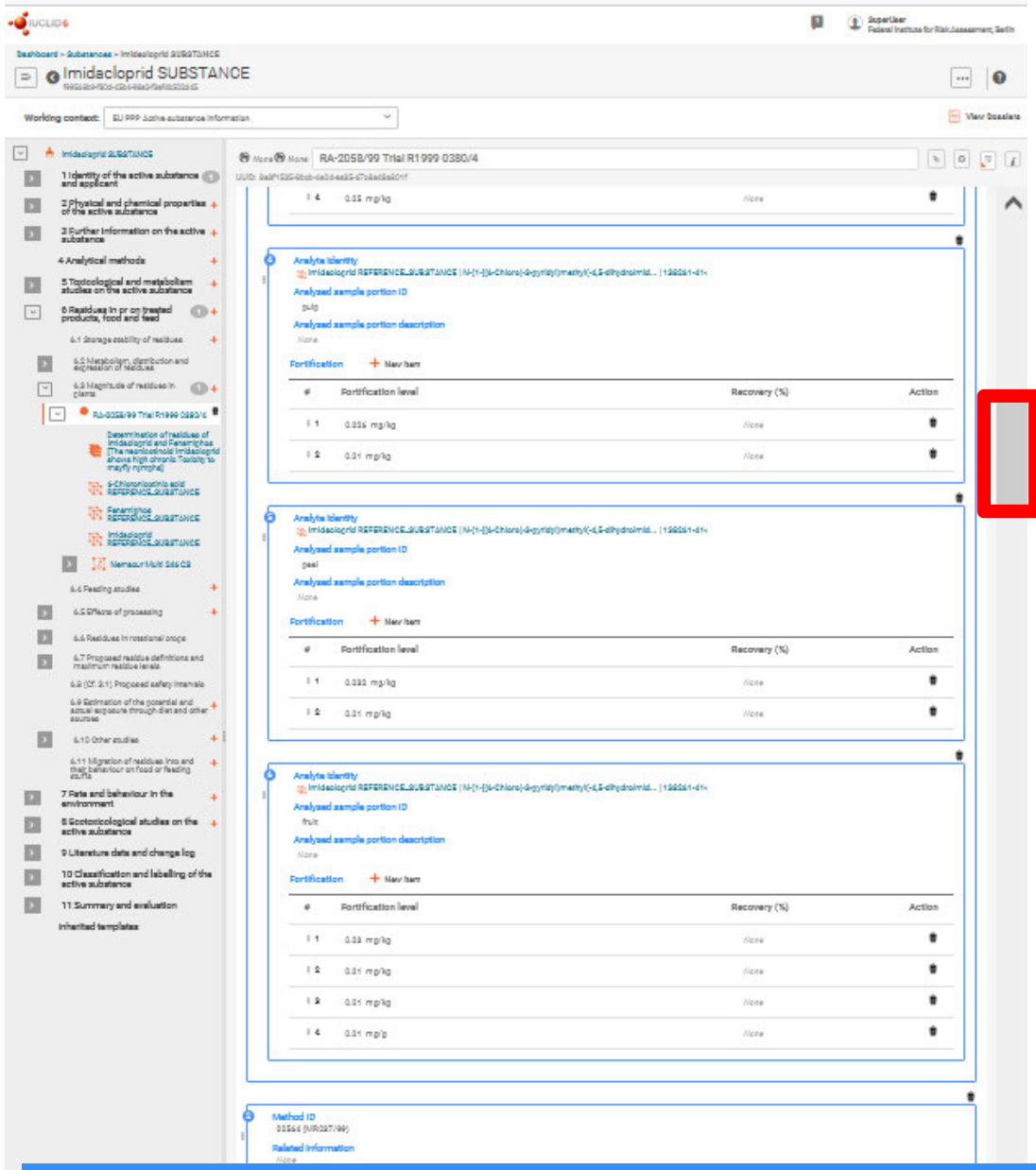
i. be based on the needs of the reviewer and not the electronic technology requirements;
ii. consider how often that field will be searched and by whom (i.e., searching is easier with fixed-fields than free-text fields);
iii. consider the need for future manipulation of both text and numeric data in specific fields, e.g., extracting text blocks and/or numeric data into evaluation reports, performing statistical analyses, data mining, or other mathematical operations. For these tasks, fixed fields generally provide a greater ability than free-text fields;
iv. Consider whether (and the degree to which) old, unstructured free-text data will be migrated into a field (migration to free-text fields is easier

The third condition, the need to manipulate numeric values and to calculate with them ➔ and use them in calculations may be valid for "**Primary result data**" but not for "Aggregated Raw Data". It would not make sense to use a dossier submission transport IT system to create pivot tables in the phase of writing the "GLP Study Report".

The history of OHT 85-5 (first OECD template with Aggregated Raw Data)

- The German XML-interface XRUEDIS was developed and proposed in 2003 to feed the German database of residue results from controlled residue trials
- The idea of an attachment type XRUEDIS was not further supported by applicants
- BfR agreed to support the extension of OHT 85-5 to transport the required aggregated raw data through the mechanisms of the OHT's
- Different versions of OHT 85-5 were published and implemented in IUCLID which are **not** tested in reality
- No applicant (data sender) is ready to support this template.

18 years later this information flow is not yet used!

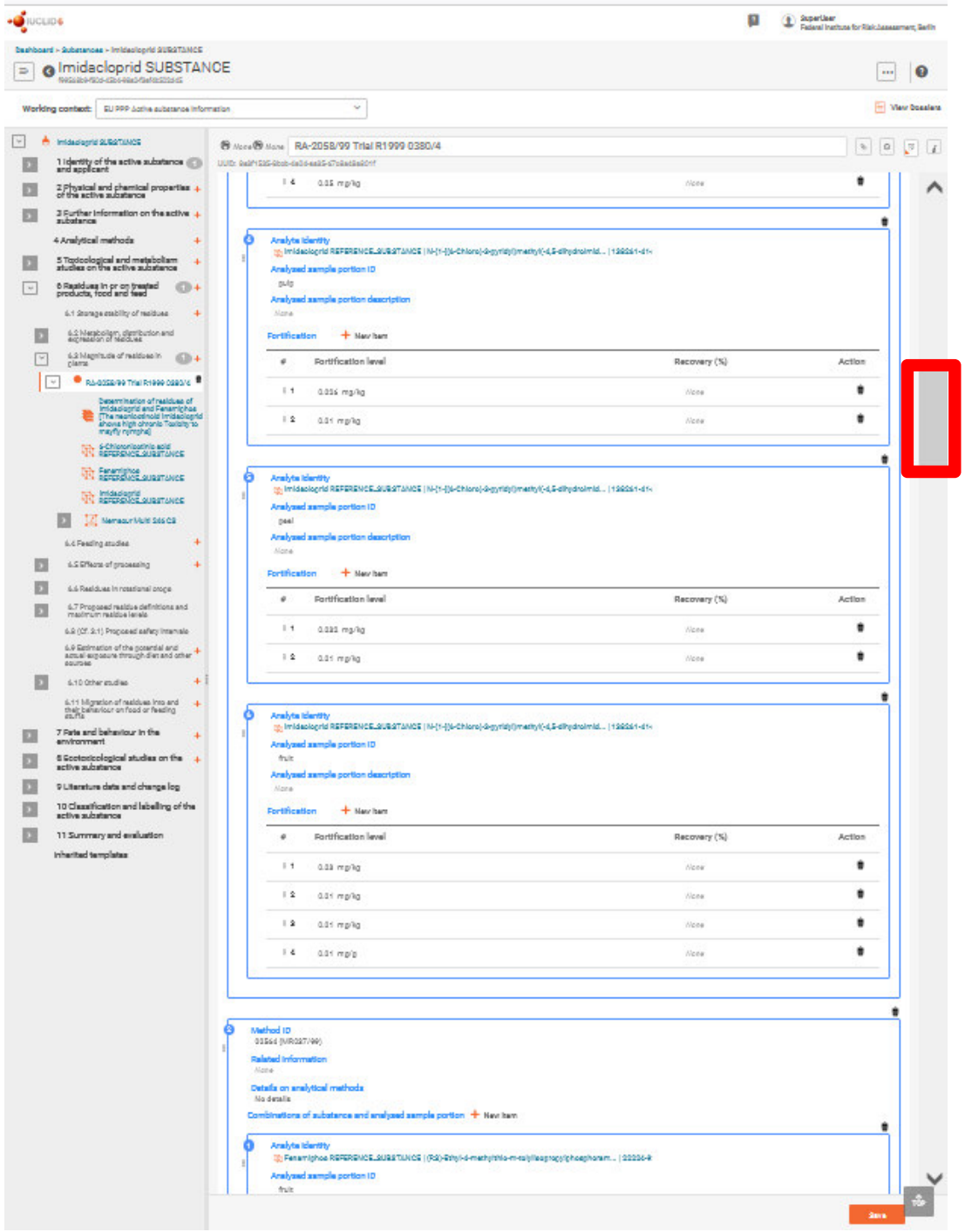


A very simple example for an endpoint study record according to OHT 85-5

- 1x Trial
- 1x Plot
- 1x Mixture
- 2x Methods
  with 2 to 4 fortifications levels
- 3x Substances
- 3x Residue values

The human readable IUCLID user interface is destroyed with aggregated raw data by the number of nested repeating block

➔ Human is not able create complex test data for programming the import tool
➔ Human is not able to understand the content and **to check for errors**
➔ Human would need “Ruedis” or an adequate internal IUCLID report which not exists



A very simple example for an endpoint study record according to OHT 85-5

Application of a mixture consisting of two active ingredients of:

Imidacloprid Fenamiphos

Analysis of the two active ingredients + 1x metabolite

6-chloronicotinic acid

Endpoint_Study_Record was created at the “Imidacloprid” substance data ➔ **OK**

This study could not be found for the “Fenamiphos” substance data set ➔ **WRONG**

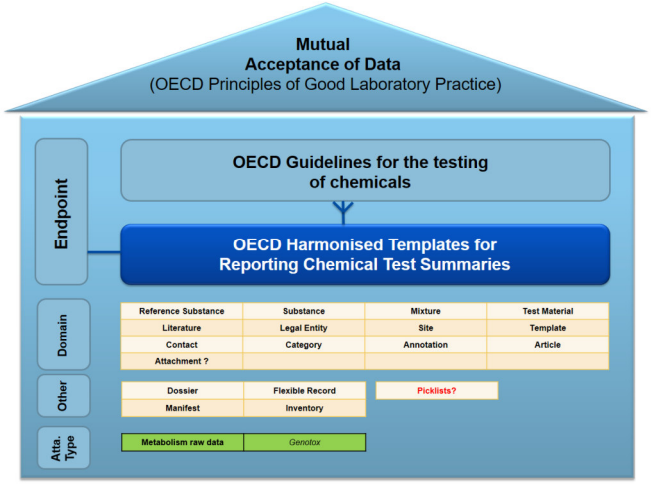
The result: A lot of disadvantages were implemented into OHT 85-5

- Only export tools would be able to create such endpoint summary records in the future
- The human readable IUCLID user interface is destroyed
  Human is not able to understand the content and to check for errors
- Only reports would be able to interpret / condense this bulk of raw data into a human readable format and no realistic scenario could be seen to get such a IUCLID report
- The internal references for residue data of mixture with 2nd additional active ingredients are implemented inadequate
- A publication of such aggregated raw data would make no additional benefit for the public

The need to submit aggregated raw data exists for residue data, as well as for metabolism studies which is in focus of this report.

The difference is, that the generic “OECD Domain Type” for metabolism raw data would be used in many harmonized templates. So, there is the realistic risk to destroy the IUCLID user interfaces for all of these harmonized templates.

The **generic** solution “Create a new OECD Attachment Type” which is recommended by the BfR



- For each **aggregated raw data type** define a specific attachment type.
- Let’s start with the metabolism study type
- **Aggregated raw data** are ONLY semantic duplicates of the data summarized in the human readable compilations of the “Study Summary Metadata” and they are not masked in word processing tables, they are submitted as separate field values

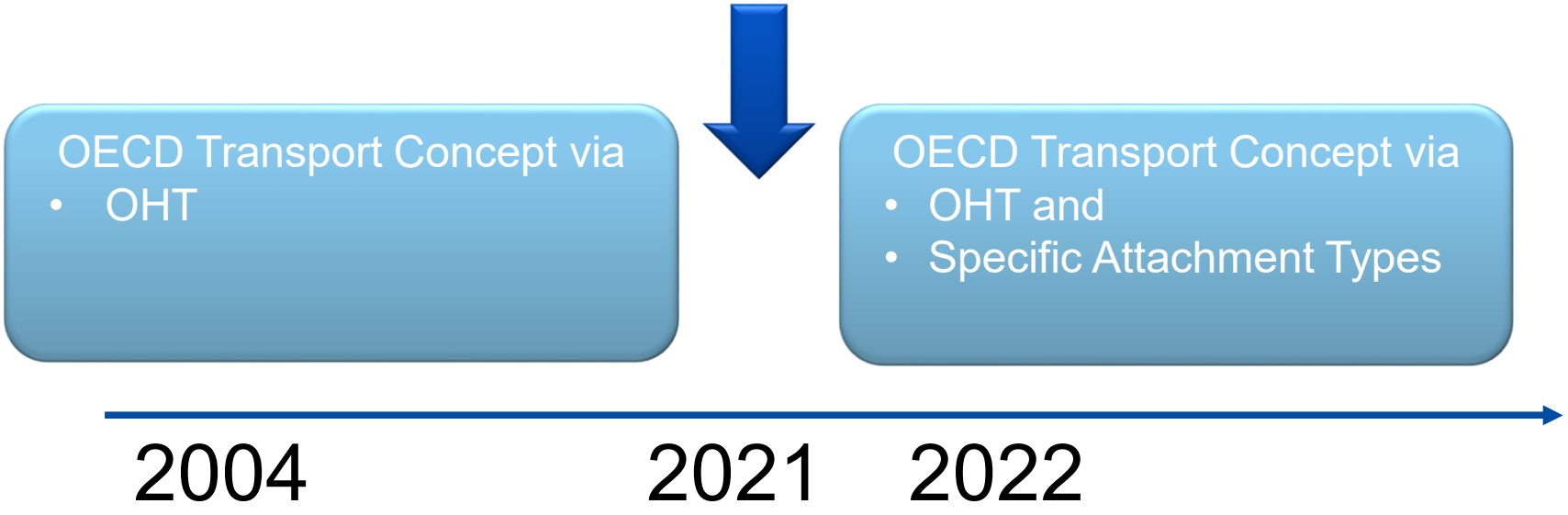
This requirement does not affect OHTs: The OECD should continue to maintain OHTs in accordance with the 2004 compromise concept, with the requirement that content transmitted via OHTs be directed to human as the recipient.

- BfR sees also a need to transport “Aggregated Raw Data” for other endpoints, e.g. “Genetic toxicity in vitro” (OHT 70)

Let’s start to define XML schema to transport these data on attachment level.

It’s time for an improvement of the OECD Transport Concept

Identification of weaknesses and handicaps in the transmission of aggregated raw data via OHT’s



- The OECD should be open for improvements according the PDCA Cycle of ISO 9001 “Quality management systems”
- OECD Attachment types would be an extension of the existing transport concept
- There wouldn’t be an impact on the IUCLID user interface and on the internal logic

Switch to the voting system now regarding the transport concept

Are there any questions? Please use the hand raise in the TEAMS environment.
For statements you could use also the TEAMS chat. The chat will be recorded. **So no idea is lost.**

Which transport concept would you recommend regardless of the political hurdles / necessary decisions?

Which functions should support the MetabolAS systems interface? (multiple choice)

GERMAN FEDERAL INSTITUTE FOR RISK ASSESSMENT

BfR
Bundesinstitut für Risikobewertung

Thank you for your attention

Stephan Worseck

German Federal Institute for Risk Assessment
Max-Dohrn-Straße 8-10 ● 10589 Berlin, GERMANY
Phone +49 30 - 184 12 - 0 ● Fax +49 30 - 184 12 – 99 0 99
bfr@bfr.bund.de ● www.bfr.bund.de/en