

Report on the MetaPath User Group Workshop regarding: “Improvements to the information flow of pesticide related metabolism studies”

Held on

10.11.2021

18.11.2021

01.12.2021

as Web-conferences

Organized by

US EPA and

BfR

Workshop report compiled by:

Stephan Wörseck

Tobias Opialla

Falko Frenzel

Table of Content

Introduction.....	2
Session 1: 10.11.2021	3
Objectives for the further development	3
Current status of the project and the aim of the workshops	3
The MetabolAS ecosystem.....	4
User requirements and concepts I.....	5
Session 2: 18.11.2021	7
Recap of the first session results	7
User requirements and concepts II.....	8
Improving the OECD Transport Concept!	10
Preferred transport option for the needed information flow	12
Needed framework conditions for laboratories and applicants.....	12
Needed framework conditions for evaluators	13
Session 3: 01.12.2021	15
Recap of the results of the 2 nd session	15
How to organise an OpenSource project	17
High-level statements.....	18
Executive summary.....	19
Establish an adequate improvement process.....	19
Closing the workshop.....	21

Introduction

According to the specific agreement under the framework partnership agreement No GP/EFSA/AMU/2020/02, proposals for the improvement of the current information flow of metabolism studies should be developed by BfR.

BfR had organised a non-formal¹ commenting procedure on the basis of a draft report version in September 2021². This draft report was circulated in the following groups:

- PSN IUCLID group
- EU Member States (MRL review) group
- EU Member States (peer review) group
- MetaPath user group (MUG)
- Pesticide companies
- Independent laboratories
- EFSA, ECHA, OECD

BfR and the US EPA had organised three MUG Meetings in November and December 2021 to discuss open questions

- regarding user requirements,
- to compare technical solutions and
- to highlight organisational questions.

The aim of this report is to document the discussion process. The participants started from very different positions, various knowledge and with different interpretations of technical words. It is thus understandable that misunderstandings were identified in the discussion as well.

The concept of the workshops was based on the assumption, that the participants had read at least the most important passages of the Draft Report published in September 2021. However, it turned out that more time was needed for discussion than anticipated.

At several points during the web sessions, participants had the opportunity to anonymously select from prepared answer options for the most relevant open questions. Only those participants who felt they were able to vote for the particular question, were asked to vote.

In the first session, it quickly became obvious, that there was a need for further discussion. For this reason, a third web conference was organised and the agendas were adjusted accordingly.

The type of discussion could be summarised by

- Analysis of the needs of the assessment process for metabolism studies which will be the basis for the next two decades
- Discussion on technical questions.
The IT support should follow the needs of the end user and not vice versa.
Selection of an optimal IT framework.
- Organisational questions to initiate and move forward the project.

Disclaimer:

It should be noted that the illustrations represent the status of the respective workshop. All contributions to the discussion were anonymised and assigned to the respective organisations. They should be seen only as contributions to the discussion in the context of this workshop.

Acknowledgement:

The presentations attached to this report have been approved for publication by the authors / organisations.

¹ outside of a regulated procedure

² https://www.bfr.bund.de/en/analysis_of_the_information_flow_in_metabolism_studies_on_pesticides-272198.html

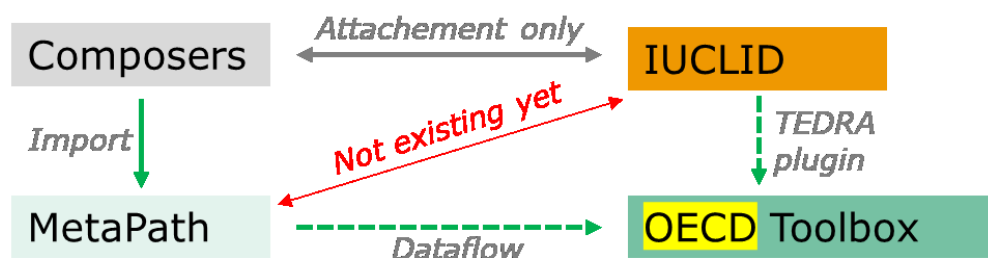
Session 1: 10.11.2021

No	Time Paris	Topic	Who
1	14:00	Welcoming Words	US EPA BfR
2	14:10	Objectives for the further development	EFSA
3	14:25	Current status of the project and the aim of the workshops	BfR Stephan Wörseck
4		Feedback discussion of main topics	
4.1	14:45	The MetabolAS³ ecosystem	BfR Stephan Wörseck
	15:35	Break	
4.2	15:45	User requirements and concepts I	BfR Falko Frenzel
5	16:45	Closing the day	US EPA

Objectives for the further development

EFSA has presented the

- framework where this project was embedded,
- the current situation
 - duplication of work to feed Composer XML files and IUCLID
 - the lack of data centralisation of metabolic pathway information
- the objectives for further development.

Discussion:

ECHA explained that ECHA is not yet actively working on an integration of MetaPath and IUCLID. ECHA is waiting for the outcome of this BfR analysis.

Current status of the project and the aim of the workshops

BfR has presented

- the status of the project,
- an overview regarding the results of the commenting phase:
 - all “important” and “very important” open questions have a technical background and make reference to the chapter 6 “Solution approaches”
 - no open questions regarding the high level summaries or the defined terms

Discussion:

- ECHA has the impression, that the BfR proposal for a general redesign with a new tool will not solve the problems. New tools might create new problems. The differences / reasons for this radical BfR proposal should be made clear.
- ECHA/EFSA also asked for a list of the user requirements with clarifications on which ones are new (not possible with the current tools) and which ones are already feasible with the current tools. This list is needed to take a decision on whether a brand new tool is needed or could be built from the existing.

³ The preliminary project name “MetabolAS” was changed to “Metapath II” at the end of the workshops (see page 19).

The MetabolAS ecosystem

BfR presented:

- the parts of the proposed new ecosystem and
- the process and objects of the information flow



Discussion:

- UK regrets that the OECD no longer supported the Metapath activities. The reasons should be analysed and OECD should revise the prioritisation.
UK wants to contact the OECD directly in relation to the current actions of EFSA (and BfR). It was proposed to request an additional agenda item to the Chemicals and Biotechnology Committee (CBC) for the February meeting.
- ECHA: A restart should not duplicate existing harmonised systems. A restart of an OECD project will need some time.

Voting:

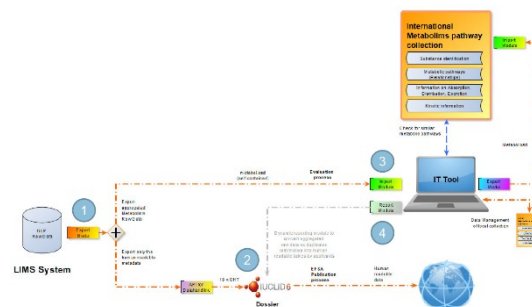
After the discussion, the participants were invited to vote with the following results:

- ~95% of the votes agreed that OECD should hold the role of the governance body for the proposed MetabolAS ecosystem.
- 100% of the votes support the implementation of an international curated reference collection of metabolism studies
- The voting regarding the question “Should the international curated reference collection of metabolism study metadata be publicly available?” shows, that all participants agree with a publication but additional discussion is needed
 - 58% of the votes: “In principle yes”
 - 42% of the votes: “Rules for the public access should be defined”
 - 0% of the votes: “In principle no”

BfR presented the attachment type as the preferred transport option

Discussion:

- What is the benefit in creating an attachment file instead of MetaPath with another newer tool?
BfR:
 - The metabol.xsd schema is needed independent from IUCLID.
This is the data interface between different MetabolAS instances.
 - With the attachment type, the MetabolAS Tool does not need additional import data interfaces.
 - Currently different composer schemes are used. This will be simplified.
 - The generic approach opens the possibility to expand these methods on other metabolism study types.
- IUCLID cannot analyse attachments.
BfR: This is OK for today but a little additional effort would be needed to enable IUCLID to ingest the necessary information from the attached file(s).
- Which part should be published?



BfR:

- All data, which are human readable with a browser application or a pdf viewer, are publishable. If an XML file contains only semantic duplicates of the provided information, there would be no information gain for the public.
- The public should read the published data according to the implemented publication rules of EFSA.
However, it is not problematic if these XML attachments would also be published after the questions of confidentiality have been answered.
- Why do we need “aggregated raw data” for metabolism studies?
 - BfR: The aggregated raw data is also needed for the residue trials and perhaps also for additional OHTs in the future e.g. for geno-toxicity data.
The users’ needs will determine, how to further process the data (summation, recalculation of values based on other reference e.g. metabolite concentration calculated as parent substance) or if these data could be included in QSAR models. Evaluators have to answer the question: Do we need this level of isolated values for our assessment procedures? If not, the technique to read the data in the OHT in the applicant’s summary Rich-Text field would be enough.
 - UK: The level of detail should support a read across and a linking to other information systems to avoid animal tests.
 - US EPA: The possibility to cover all other metabolism guideline study types would be enhanced with the new system. Evaluators need metabolic simulators.

Voting:

After the discussion, the participants were invited to vote for the question: “Which transport concept would you recommend regardless of the political hurdles / necessary decisions?” The result shows, that all participants prefer an OECD solution and that an additional discussion is needed:

- 53% of the votes: “Create an OECD Domain Type”
- 47% of the votes: “Create an OECD Attachment Type”
- 0% of the votes: “Use of 3rd Party Attachment Types”

ECHA pointed out that not all possible options were available for selection on this question. Therefore, some participants chose an option which was closest to their preference. However, this does not mean that they were convinced by this option.

User requirements and concepts I

BfR has presented

- different chemical structure notations,
- the need of generic Markush notations

Discussion:

- Which chemical notation would be better than SMILES?
The Metapath and the (Q)SAR toolbox are using the SMILES notation.
- Possibility of conversion of SIMILES into InChI and vice versa?
Is it possible to migrate MetaPath SMILES codes in the MetabolAS InChI code?
Regarding the generic Markush notation, LMC noted that they have - in a customer version of MetaPath - provided their own concept of generic notation based on SMILES. This function could be provided next year in a general MetaPath update⁴.

⁴ Done with MetaPath ver.5.4.0 and MSS Composers ver.1.10

https://oasis-lmc.org/downloads/documents/Whats_new_MetaPath_Composers_2022_January.pdf

Voting:

After the discussion, the participants were invited to vote for the relevance of the need of a generic notation:

- 40% of the voters agreed, that generic Markush notation is needed to a certain extent in metabolism studies.
This means, that coding of generic structures is a “show stopper” for this project.

BfR has presented

- different chemical structure notations,
- the need of generic Markush notations,
- the need of grouping and calculations and
- methods for visualisations

Discussion:

- EFSA: Why do we need grouping and calculations in MetabolAS when it is already implemented in Ruedis?
BfR: Ruedis manages residue data, not the metadata of metabolism studies. The reference to Ruedis during the presentation was only to show which user functions could be prompted to the evaluators and provide a benefit in the assessment process.
- EFSA: All the existing good functions of MetaPath should be implemented/adopted with priority.

Voting:

After the discussion, the participants were invited to vote for the relevance of grouping and calculation in the assessment process:

- 75% of the voters agreed, that evaluators need user functions for grouping and calculation in the assessment process.
To provide such functionality, the MetabolAS has to manage the isolated reported measured values instead of reporting them in HTML tables.
- Two participants noted in the voting systems, that they have additional requirements for visualisation. Unfortunately, the ideas were not further discussed in detail.

Session 2: 18.11.2021

No	Time Paris	Topic	Who
6	14:00	Greetings / welcome to the 2nd day	US EPA
7	14:05	Recap of the results of the first session	BfR
8		Continue "Feedback discussion of main topics"	
8.1	14:15	User requirements and concepts II	BfR Stephan Wörseck
9		The OECD transport concept for "Aggregated raw data"	
9.1	14:50	Improve OECD Transport Concept!	BfR Stephan Wörseck
	15:15	Break	
9.2	15:25	Proposed: Preferred transport option for the needed information flow	ECHA
10		Needed framework conditions	
10.1	15:55	for laboratories and applicants	CropLife Europe
10.2	16:25	for evaluators	EFSA Anja Friel
			ANSES Gaelle Vial
			BfR Thomas Kuhl
			US EPA
11	16:55	Closing the workshop	US EPA

Recap of the first session results**Clarification of the role**

EFSA asked for a clarification of the roles in this improvement process:

- BfR's role is to make proposals based on the in-depth analysis of the current framework conditions and the needs of the evaluation process.
- The Metapath User Group should help to clarify open questions, to support BfR proposals or deprioritise aspects if they are not needed.
The MUG is representing the adequate user forum. If a solution is later to be called a customer-oriented solution, then this forum of experts must be given appropriate importance.
- EFSA would be in charge of evaluating and making internal decisions in 2022. This EFSA decision-making process is the topic in the second part of the 3rd web conference. But EFSA is not in charge of taking decisions alone. Such high level decisions on the MetaPath improvement, on data transport, etc, are to be taken with all stakeholders. EFSA can facilitate the discussion and the coordination but shall not be perceived as isolated deciding authority.

Voting:

After the 1st meeting, BfR gained the impression, that it was not clearly expressed whether IUCLID should be improved to fulfil all user requests regarding the evaluation of metabolism studies.

Participants were invited to vote: "Is an improved MetaPath system (MetabolAS) needed in addition to IUCLID in future?":

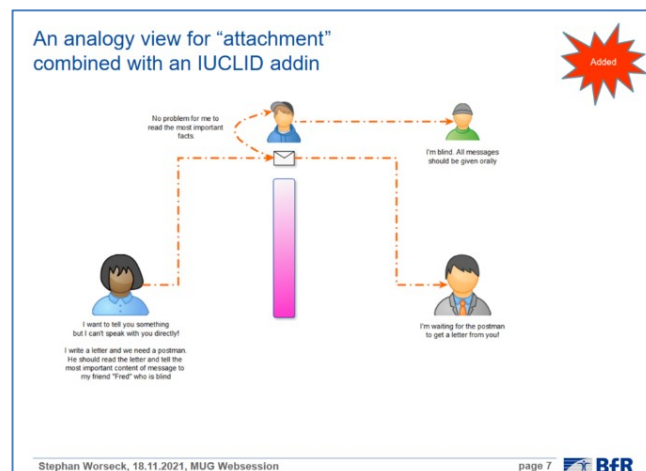
- 86% of the votes agreed that an improved MetaPath (MetabolAS) is needed as an evaluation tool as well as the backbone for the curated reference collection of metabolism studies.

The voting and the following discussion emphasised the need for an improved MetaPath (MetabolAS):

- ECHA noted, that there should be a clear separation of processes and no duplication of functionality of both systems.
- EFSA noted, that there is a real need for an improved MetaPath for risk assessment.

BfR has presented an analogy for the transport terms “attachment” and “OHT”.

BfR proposed to combine the attachment type and the need of IUCLID to harvest some data regarding the substance list and the relation between the substances via an plugin into the IUCLID system. In this case, a communication without any transformation could take place on attachment level and IUCLID would be able to consume the relevant data. However, no participant had given a comment to this proposed combined transport model.



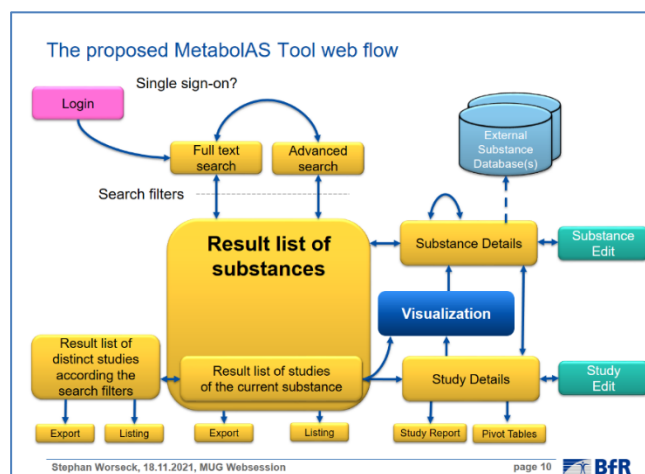
User requirements and concepts II

BfR has presented additional user requirements in addition to the first session regarding the:

- Database
- MetabolAS Tool
- Data Management in the Reference Collection
- Management of (Q)SAR results
- Migration

From BfR's point of view, it is considered that the MetaPath application and the used database implementation have serious weaknesses which culminates in a requirement for general redesign. The XML schema of the MSS/DER composer family is appropriate to transport the raw data values to the MetaPath system. This data model would be open for a migration to an improved MetaPath (MetabolAS) system.

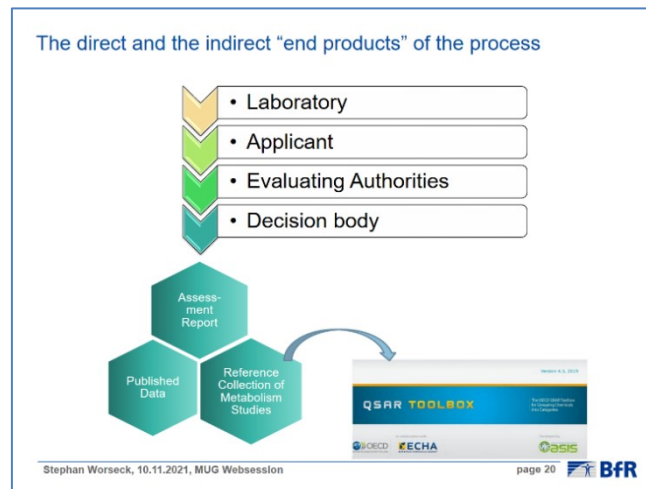
The proposed MetabolAS Tool could be the same tool on applicants and authority side because of the need of the same functionalities. This would be the same situation as today with MetaPath.



Discussion:

- From ECHA's point of view, the proposed MetabolAS Tool is a repetition of functions and technology of IUCLID on basis of the data of the OHT's. IUCLID should be used as the data collecting system after the GLP report is ready.
- At the beginning of this session EFSA noted, that there is a real need for an improved MetaPath for risk assessment.
- Although a project meeting was held between EFSA, ECHA and BfR, ECHA was not able to present its concept on how to integrate the different tools without replication in a comprehensible way. The reference that the ECHA concept is based on already published international standards was not sufficient for BfR.

- It seems that BfR was not able to communicate to ECHA /EFSA the difference between showing the “submitted data” and
 - to work / recalculate / aggregate / compare / weight ... with these data
 - to create the evaluated data set as a “direct end product” of the assessment process which has to feed the curated reference collection of metabolism studies and via the reference collection (Q)SAR Tools indirectly
- ECHA recommended that the necessary functions of MetaPath (MetabolAS) should be aligned with the existing functions in IUCLID and preferred the motto:
“If one of the functions is already implemented, let us use this tool”.
- In ECHA's opinion these decisions should be taken by this user group (MUG) or even better by the OECD because they can ensure that the governance is stable in the long run.
- EFSA emphasised, “We need interoperability to share data - not necessarily new system development.”



Voting:

There was one open question regarding the API: “Is it realistic / purposeful to plan a data transfer from a LIMS systems into a local MetabolAS collection via APIs and / or common used access methods ?” Only applicants were asked to vote.

- All of the four votes agreed that these functions are needed.

The voting for the question: “How often do you have aggregated raw data in a simple electronic format (outside of LIMS)?” should help to prioritise the functions to import alternative formats.

Voting result:

- The amount of other digital data outside of LIMS is relevant to support a flexible alternative import function.

“Would you say a bulk import of substances / dose groups or result tables is possible most of the time?”

Voting result:

- Six of eight voters agreed.

“If a bulk import of files would be possible, is an interpretation of tables via the clipboard still necessary?”

Voting result:

- Four of seven voters agreed.

“Are there additional phys-chem properties (beside $\log P_{OW}$) or toxicological data as structured meta data needed for the evaluation process?”

Voting result:

- Three of five voters agreed.
It was not discussed, what the parameters might be!

“How often do you get late additional structure (or name) information on initially “unknown” metabolites reported in the original GLP report?”.

Voting result:

- There were seven votes.
Five of seven participants said, this happens for “5% to 20% of the reported metabolites”. So this is an important fact which should be reflected in the user requirements and the use cases.

“Should the MetabolAS Tool be able to manage (Q)SAR responses for each substance from different (Q)SAR Tools according the ECHA guide?”.

Voting result:

- There were 11 votes.
1/3 of the participants voted for such a functionality.
It seems this is undecided but has no priority and should be discussed in the future by the experts.

BfR has presented some important points which should be solved in the migration phase:

- Matching from strings to picklist items
- Check valid dependencies between elements (e.g. Goats and rat strain)
- Split HTML tables
- Split groups which are concatenated by comma

Voting:

The effort for migration is estimated to be 20% of the whole project. “Should the migration start from the XML file?”

Voting result:

- There were 6 votes with two stating that it’s better to start from MetaPath.
Only one participant had given a reason.

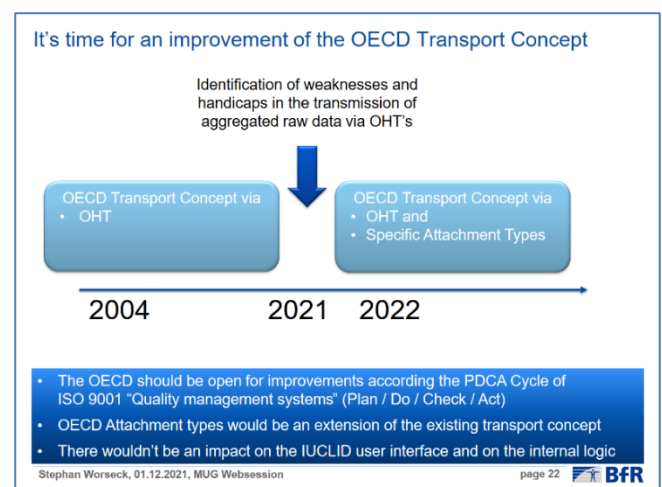
Discussion:

- LMC makes databases mainly with the MetaPath interface. They have collected 1500 maps. LMC proposed to convert the whole database to XML files for a migration or to start the migration from the MTB database.
But LMC was the only user who has spoken about such a process. The other of the two participants remained anonymous until the end.
- A deeper analysis is required to decide from which data source the migration should start.

Improving the OECD Transport Concept!

BfR has presented its position regarding the question of the right balance between the transmission of information in several individual fields and in aggregated texts is very old and goes back to the founding days of the OECD Harmonised Templates.

- The final Report of the Expert Group Meeting to Explore Harmonising Templates ENV/JM/RD(2004)9 gave principles to find such a good balance. The



first principle was: “be based on the needs of the reviewer and not the electronic technology requirements”. This principle was neglected more and more.

- The current OHT 85-5, the first Harmonised templates where aggregated raw data are included, has shown:
 - Humans are not able to create complex test data for programming the import tool manually
 - Humans are not able to understand the content and to check for errors
 - Humans would need “Ruedis” or an adequate internal IUCLID report tool, which shall be in development at the moment.

The need to submit aggregated raw data exists for residue data, as well as for metabolism studies, which is the focus of this report. BfR sees the need to transport “Aggregated Raw Data” for other endpoints, e.g. “Genetic toxicity in vitro” (OHT 70) as well.

Because the generic “OECD Domain Type” for metabolism raw data could be used in many harmonised templates (~18 OHTs), there is the realistic risk to destroy the IUCLID user interfaces for all of these harmonised templates.

The human readable IUCLID user interface fails with aggregated raw data due to the increasing number of nested repeating blocks. Humans will not be able to understand the content and to check for errors. Humans would need an adequate internal IUCLID report to translate the “Aggregated Raw Data” back to human readability.

Discussion:

- EFSA noted that, the manual data input of aggregated raw data is possible in the MSS-Composers. This is a very heavy workload and two days per study are needed. If this is to be done by IUCLID, the granularity would be the same. In addition, there would be the chance to collect these data from the beginning of the process. Moreover, EFSA prefers to use this chance in the European PPP processes. The question is: Why is it more difficult to expand the OHT in IUCLID than in the composers?
- BfR: The granularity of aggregated raw data is not the same. The MSS composer using only HTML tables where the values are included in table cells. This format is **not** adequate for calculating, recalculating or grouping values. So one would have to input isolated values and for each value all the references to the object of investigation, the dose group, the sample group, the method etc.
So this would increase the data input by a factor of 3 to 5. There is no known need for this level of detail in the human user interface of IUCLID.
- EFSA noted that the information should be included in IUCLID according to transparency regulation.
- BfR: That would be realised because a semantically identical information will be submitted on attachment level (for machines) **and** in the human readable format in the applicants summary.
- This is the reason, why BfR proposed the improvement of the OECD transport concept. The international processes need much more aggregated raw data, which should lead to an enhanced OECD transport process.
- EFSA noted that this group has developed a good MetaPath application, which is helpful in the evaluation process. If we are creating new tools, we need time to validate those tools to get them accepted and adopted. So one should just think about how MetaPath and the MSS composers can be improved, particularly on the key areas that people would like to see improvement on.
- BfR added a note: The EFSA’s opinion shows that EFSA is not yet prepared to accept the general need for a reform of MetaPath and the MSS composers. Within a

software application lifecycle, one regularly reaches the point where technological jumps are necessary that are comparable in scope to a new development of the software. In some cases, maintaining backwards compatibility is even a real cost factor. If these technological "jumps" were not enacted, we would still be working with Word97 under MS DOS today.

- ECHA noted that the OECD transport mechanisms in this area have been under continuous development since 2004 and that the OECD is open for discussion.
- BfR noted, that a request to the OECD only makes sense if this is supported by
 - the MetaPath User Group,
 - by Germany or even better
 - by Europe as a whole.
- US EPA added additional aspects:
 - There are not only bugs in the current systems. Users also want to have additional functions for the risk assessors, which would make their jobs easier.
 - The current concept of MSS composer is not able to capture other guidelines with radio isotopes.
 - Who should host the new service?
 - One should not look only to the resource costs in terms of money but also on the time requirements. Are we looking at 5 or 10 years?

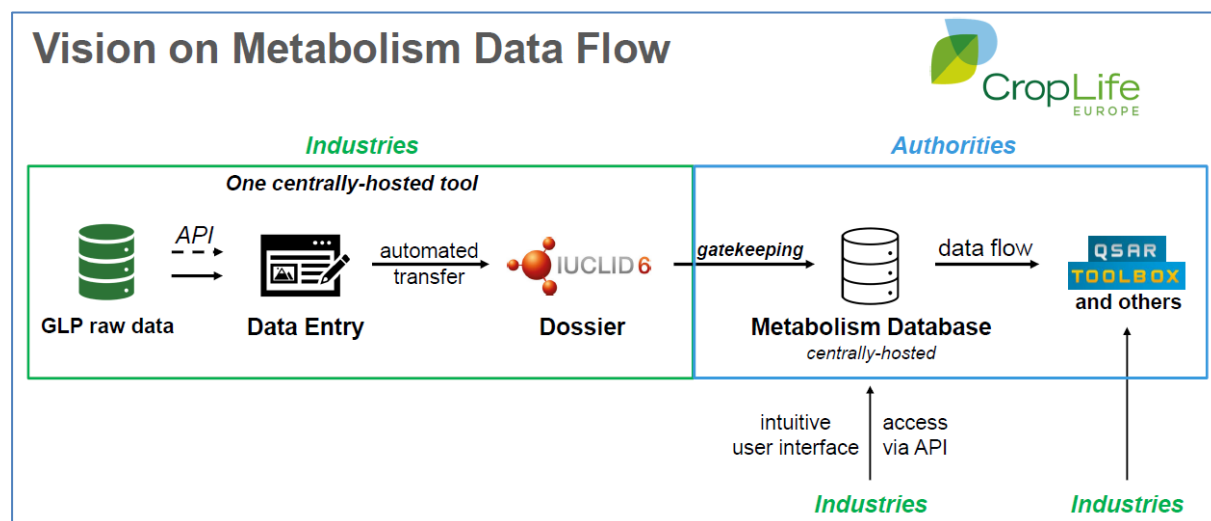
Preferred transport option for the needed information flow

ECHA had no separate presentation to this topic. The ECHA position was already clearly explained : "Our proposal is to reuse systems and not to create new tools."

Needed framework conditions for laboratories and applicants

CropLife Europe presented their point of view regarding the metabolism dataflow:

- the current weaknesses of MetaPath and the four Composers which are centrally managed but not published in a transparent way e.g. with sending a newsletter to subscribed users.
- the most important critical point is the user-friendliness, and the poor performance of these tools
- the need for an API to MetaPath and the Composers
- the need for automated transfer of the collected data to IUCLID
- the need of a metabolism database which is centrally hosted



The improvement of the metabolism dataflow will be one of the most important issues, if this project would be done in a harmonised and long-term approach on a global scale. Drastic improvements and rearrangements are needed to really meet all the requirements from the applicant's side as a minimum..

Discussion:

- EFSA asked about the confidentiality aspects of a centrally curated reference collection of metabolism studies.
CropLife Europe: Gatekeeping and data protection during the evaluation process is very important to industry. It should be possible to access a global reference collection of metabolism studies as well as creating local instances. The used software should always be up-to-date.
- BfR asked regarding the needed functions in the assessment process on the applicant's side.
CropLife Europe: The development of an active substance is an ongoing process on industry's side. If an assessment is done at the end, one really has all data together.
- EFSA noted several times that the improvement process should take 2 years at maximum, not 5 or even 10 years. Therefore, EFSA is not interested in a radical redesign.
- There were different opinions regarding the need for uniqueness of metabolism studies in the database collections. The BfR prefers the uniqueness solution in order to enforce the rule "One study - one assessment". ECHA prefers mechanisms for the subsequent detection of duplicates (deduplication).

Needed framework conditions for evaluators

EFSA presented the risk assessor's view regarding the metabolism dataflow.

- Improvement should be in line with the upcoming assessment demands on OECD and EU level.
- Publication of the new OECD guidance is planned for 2023.
MetaPath has to be integrated with other tools (e.g. OECD Toolbox) to facilitate assessment of pesticide metabolites.
- The needed functions for risk assessors are:
 - Support & facilitate
 - Prediction
 - Extend & connect
- Only with a tool like MetaPath, will it be possible to fulfil the new OECD guidance
- Support the general principal also beyond pesticides area: 'One substance – One assessment'
- To initiate a "read across approach" in other tools would be the most important new additionally needed function in MetaPath at the substance level.

ANSES presented some aspects for improving the existing functionalities of MetaPath:

- ANSES has implemented the usage of MetaPath into the evaluation process.
- Grouping of metabolites and read across are important functions that ANSES has been waiting for many years.
- A central world wide database of metabolism studies and for the residue definition would be helpful.
- Support the colleague from US EPA in their view, that small local databases are needed for specific projects.
- ANSES supports the view that a lot more maps are needed.
- One should not risk losing any metabolic pathways which have been entered during the last 10 years!

- ANSES asked to keep the name [MetaPath], the proposed one has ambiguities in French.
- What has to be kept:
 - Maps and metabolites visualisation
 - Comparison of maps
 - Identification of common metabolites, search tools
 - All the work already done to populate the database
- What has to be added/improved
 - Handling few bugs
 - Improve the possibility / the way to use data feeding in the MSS Composers?
 - editing reports
 - Extraction of data from metabolism studies to generate summary tables, list of metabolites
 - Prediction
 - Kinetics

BfR presented some aspects for improving the existing functionalities of MetaPath

- In house, BfR has a lot of knowledge data bases of qualitative and quantitative data like residue data (Ruedis), Exposure, Environmental monitoring data etc. but it is not possible to combine these data with MetaPath
- A better integration to these existing information bases is needed. This should be understood as “Interoperability” of MetaPath.
A key is to combine quantitative data with qualitative data

ECHA had given a nice short summary of the 2nd session:

“There is a large agreement on the requirements, not necessarily on the solution, but that is why we are discussing today!”

Session 3: 01.12.2021

No	Time	Topic	Who
	Paris		
12	14:00	Greetings / welcome to the 3rd day	US EPA
13	14:05	Recap of the results of the 2nd session	BfR
14		Organization of a project	
14.1	14:15	Possible project management workflow	BfR Stephan Wörseck
14.2	14:45	How to organise an OpenSource project	BfR Tobias Opialla
	15:05	Break	
14.3	15:15	High-level statements	BfR Stephan Wörseck
14.4	15:45	Executive summary	BfR Philip Marx-Stöltzing
15	16:00	Establish an adequate improvement process	EFSA
16	16:55	Closing the workshop How to publish a workshop report?	US EPA

Recap of the results of the 2nd session

BfR presented a clarification of the term “Aggregated raw data” to explain the different transport levels which are needed in the information flow of metabolism studies.

“Aggregated raw data” are suitable for calculations and for grouping of results. According to BfR, this level of detail makes no sense for the harmonised templates but make sense on the attachment level. Apart the MetaPath system, there is no other IT tool known that requires this data in this level of detail.

According to BfR, the output of the flexible pivot reports of MetaPath should be included into the human readable Rich-Text fields of the OHT's.

What does it mean: Aggregated raw data ⇔ Textual summary?

Object Group	Object	Substance	Dose Group	Sample Group	Time	Value	Unit
1 Crop	Sugar beet	Parent 1	DG1	Plant	1d	4	mg/kg
3 Crop	Sugar beet	Parent 1	DG1	Plant	5d	2.5	mg/kg
4 Crop	Sugar beet	Parent 1	DG1	Plant	10d	0.8	mg/kg
5 Crop	Sugar beet	Parent 1	DG1	Plant	30d	0.1	mg/kg
6 Crop	Sugar beet	Parent 1	DG1	Plant	180d	<0.001	mg/kg
7 Crop	Sugar beet	Unknown 001	DG1	Plant	1d	0.1	mg/kg
8 Crop	Sugar beet	Unknown 001	DG1	Plant	5d	0.001	mg/kg
9 Crop	Sugar beet	Unknown 001	DG1	Plant	10d	<0.001	mg/kg
10 Crop	Sugar beet	Unknown 001	DG1	Plant	30d	<0.001	mg/kg
11 Crop	Sugar beet	Unknown 002	DG1	Plant	1d	0.1	mg/kg
12 Crop	Sugar beet	Unknown 002	DG1	Plant	5d	0.001	mg/kg
13 Crop	Sugar beet	Unknown 002	DG1	Plant	10d	<0.001	mg/kg
14 Crop	Sugar beet	Unknown 002	DG1	Plant	30d	<0.001	mg/kg
15 Crop	Sugar beet	Unknown 002	DG1	Plant	180d	<0.001	mg/kg
17 Crop	Sugar beet	M001	DG1	Plant	1d	0.5	mg/kg
18 Crop	Sugar beet	M001	DG1	Plant	5d	0.2	mg/kg

Example of the level of detail of aggregated raw data.

This level makes no sense for the OHTs. No other data consumer of this level of detail exists.

MetaPath has to support flexible pivot reports

Table ... : Average Substance concentration in sugar beets 24h after application of ... Parent 1	DG1	DG2
M001	0.50	0.60
M002	0.56	0.76
Parent 1	4.00	50.00
Unknown 001	0.10	0.50
Unknown 002	0.10	0.30
Sum	1.05	10.43

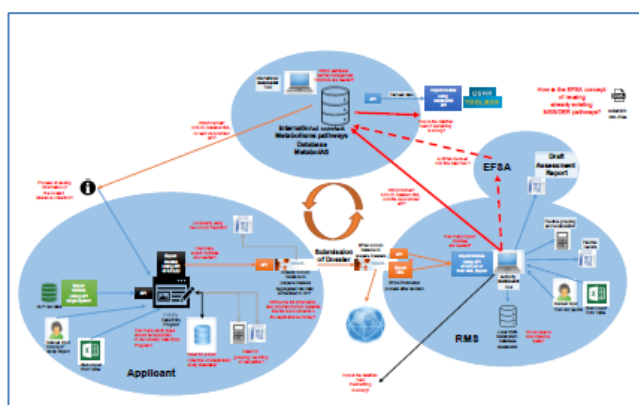
This is the level of detail needed in Rich-Text fields of the OHT's

e.g. "Applicant's summary and conclusion"

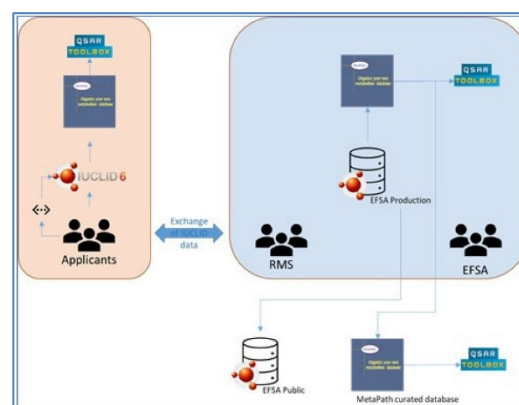
Stephan Wörseck, 01.12.2021, MUG Websession

page 3 BfR

BfR presented results of a follow up discussion between BfR and ECHA after the 2nd session to clarify the interplay of the improved MetaPath and IUCLID.



BfR scheme with questions



ECHA proposal

The BfR's interpretation of the ECHA proposal of the data flow schema:

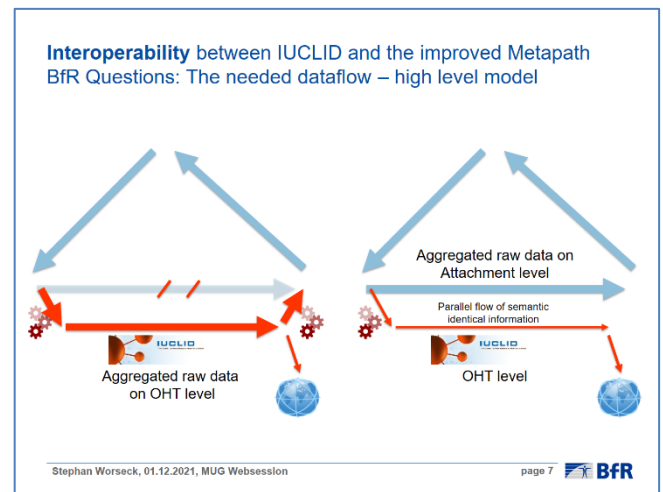
- ECHA focuses its efforts on the information flow between applicants and authorities only.

- The interface for the necessary data flow between the different MetaPath collections was not specified.
- The flow of information from the curated database back to the applicants was not considered.

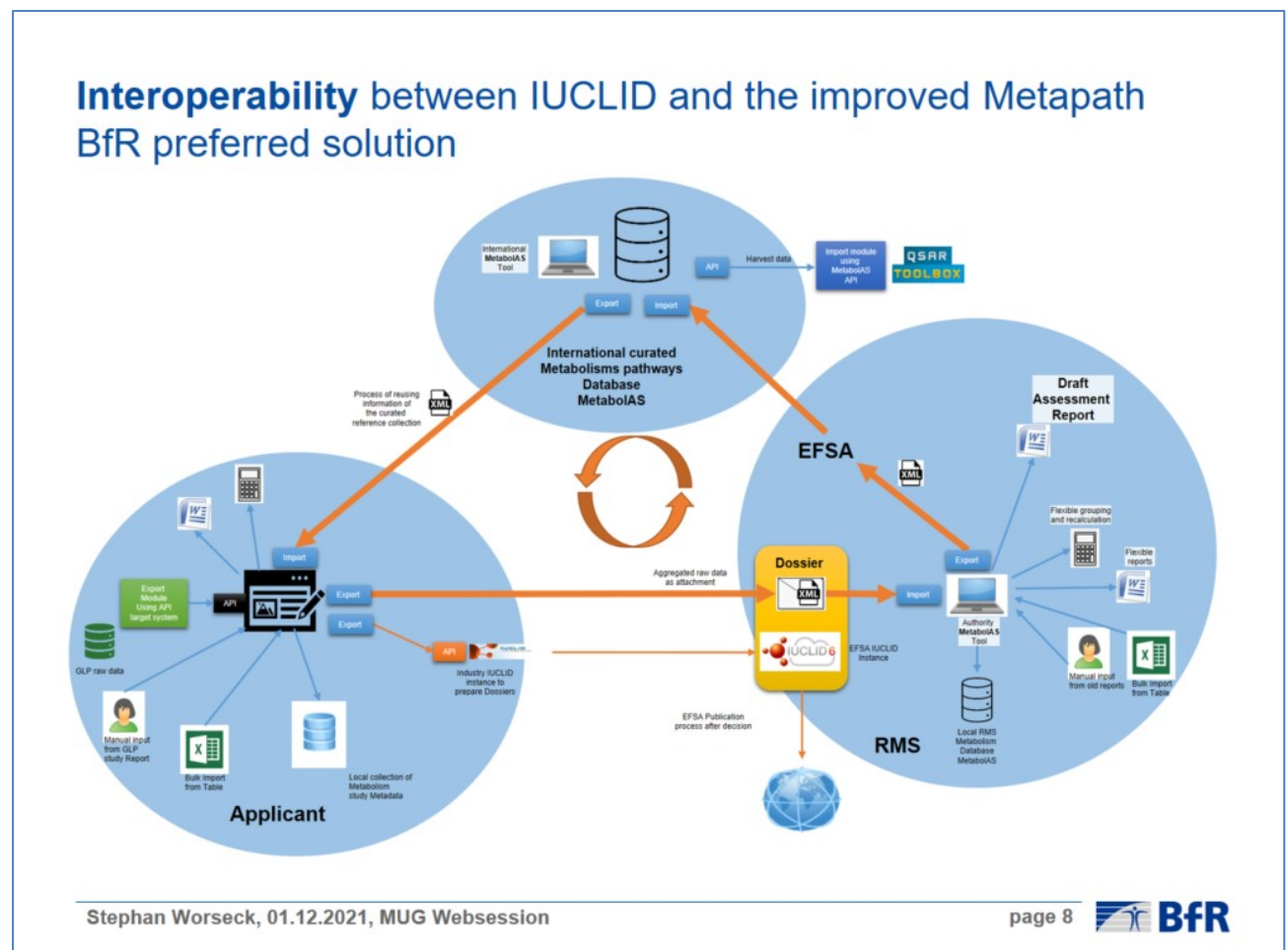
BfR presented a comparison of the consequences of the two different transport options.

- “Aggregated Raw Data” embedded in “OHT”
- “Aggregated Raw Data” as attachment

During the web session, these differences could not be further clarified, as no representative of ECHA was present due to scheduling conflict.



Based on the advantages of the BfR proposal, the following information flow was developed as the basis for the final report.



BfR has presented a decision tree, which should be agreed by EFSA and a list of adequate milestones. After that, EFSA would be able to organise a project in at least five working packages, which contain a sequence of potential project tasks for the required improvement process.

The work packages were phrased under the assumption that the programming will be organised via an “Open Source Project”. Otherwise, EFSA would need to compile the user requirements, schema definitions, the mock ups, algorithms, XML data files into a specification for a tender process.



Discussion:

- CropLife Europe support the analysis of BfR (see “[Final Statement on the Metabolism Data Flow](#)”). IUCLID could not be the place to collect the aggregated raw data of metabolism studies. A separate data entry platform would be appreciated. Contractors, who are filling the IUCLID dossiers and the composer XML files, do prefer the composer scenario. CropLife Europe prefers different data entry tools. Submission of metabolism data should be easier.

How to organise an OpenSource project

BfR presented aspects of

- the term definition of OpenSource,
- why OpenSource is necessary in regulatory frameworks
- the structure of an OpenSource project and
- how to organise it in an OpenSource ecosystem

Discussion:

- What do you think about the proportion of the efforts for project management in an OpenSource project compared with a project with an external contractor?
- BfR: It can be assumed that the part of the programming is identical in both cases. Looking at a short period of time, a closed source product is cheaper. However, in the long run open source is always less expensive.
- CropLife Europe: Industry is really supporting the idea of initiating an OpenSource project but a clear governance structure would be needed for such a project that would ideally be on a global level, so that it's not only Europe focused. In addition, very strong partners are mandatory. An OpenSource project would also be helpful to overcome the technical limitations and the available resources of LMC to develop. CropLife Europe appreciates the scientific knowledge of LMC. CropLife Europe also thinks that LMC will be engaged in an OpenSource project. LMC is an important partner for this project and it should remain this way.
- BfR: This is extremely important! Authorities are not software developing companies. So contractors are also needed in an OpenSource project. EFSA may specify that certain deliverables are to be provided in an OpenSource project.
- CropLife Europe has recently seen that the authorities are very reluctant to engage directly with industry. How can this work for such an OpenSource project?
- BfR: This OpenSource project does not concern the regulatory context and also not the legal decisions.
- CropLife Europe: The idea of an OpenSource project should be assisted by industry. However, any problems should be addressed early. If the industry invests money and resources in this project, the necessary decision-making process should be very transparent.

- LMC: Today nobody can give a statement for LMC regarding taking part in an Open-Source project. All of their software is private.

Voting:

Participants were invited to vote on the question: “Do you think that an OpenSource project could be an adequate organisational format?”

- 11 of 12 votes agreed, that an OpenSource project could be an adequate organisational format for the improvement of MetaPath.

High-level statements

BfR presented the draft high-level statements, which can be typified as follows:

- Opinions (Assessment of the current state by BfR)
- Proposal (BfR's recommendations for further development)

The high-level statements are summarising needs for / to

- Harmonisation,
- a generalised concept of the term metabolism,
- an ecosystem of necessary components,
- an appropriate transport concept of metabolism study metadata,
- a curated reference collection of metabolism study metadata,
- an improved data management and data handling procedure,
- improved reports,
- organise the improvement process in an interim period and for
- an improved authority process

BfR hopes that fact-based decisions, not political decisions, would be made. If radical changes are needed, they should be initiated. BfR is convinced that EFSA is the best place to take the leading role in this process.

Due to ANSES' comments on the proposed project name MetabolAS implying that it would contain ambiguities, BfR invited participants to vote for a better name. BfR would use the favourite in the final report.

Voting:

“Which name do you prefer when it comes to an updated/improved MetaPath?”

Project name	Votes (multiple choice)
MetaPath	5
MetaPath II	10
<i>i</i> MetaPath or MetaPath <i>i</i> (improved MetaPath)	2
<i>e</i> MetaPath (extended MetaPath)	3
<i>g</i> MetaPath (generalized MetaPath or global MetaPath)	3
OECD MetaPath (comparable to OECD Toolbox)	16
MetabolAS (Metabolism Assessment System)	2

Discussion:

- EFSA: This was not only a vote for a name. The vote very clearly showed the wish of the participants, that this project will be a project under the governance of the OECD. However, this is still not agreed and is therefore an open point.
Another point: It was agreed that a closer connection of MetaPath and the OECD (Q)SAR Toolbox should be established. Therefore, it would be preferential to have it under the same governance. But there should not be duplicate functions in MetaPath

and in the OECD (Q)SAR Toolbox.

So, this vote was not only for the name.

- US EPA argues in preference for the much more generic concept name of MetaPath II.
- BfR: The current position of the OECD is that they are following the communication in the group but they are waiting until an official project will be submitted to the OECD.

Executive summary

BfR presented the draft executive summary:

- This chapter is directly addressed to the originator of this report, to EFSA.
- BfR recommends that EFSA would be in charge of organising the decision making process for the required improvement process for the metabolism studies.
- A set of high-level decisions are needed. All relevant stakeholders should be included in this process.
- It is recommended that no technical decisions are placed on this level.
- BfR recommends EFSA to establish an adequate project structure.

Establish an adequate improvement process

EFSA has been invited to discuss the process to coordinate the software change requests regarding MetaPath and the MSS Composer Family.

Why is it so difficult?

- The needs and expectations are different for data submitter, data manager, data assessors and researcher.
- There are different types of requests (bug fix, improvement, new features, need for interoperability).
- Different jurisdictions with different end products of the assessment procedure, with different terms (North America - DER, Europe - DAR).
EFSA does not have the role to manage the short term improvement process.
- Limited resources (Funding, LMC resources).
- Cannot wait for final governance decisions.
- The development of "private" MetaPath and composer versions is not helpful. The short term improvement should result in one publicly available software which meets the needs of all users/stakeholders.
- Project organisation should ensure best use of resources without redundancy and contradictions of the modifications.

All these problems require strategic planning, alignment and prioritisation.

Discussion:

- EFSA: There is no central management of all user requirements for the short term. Without such an organisational form it is difficult to manage this. The implementation of the generic Markush-Structures has shown the complexity of this change request. There are different pre-requisites to make this a positive investment
 - The LMC algorithm should be tested before going into production.
 - Is quality control of the LMC algorithm planned?
 - It only makes sense if all components (MetaPath and the composers) are improved at the same time.
 - Database / XML Files should be migrated.
- BASF: The usability of generic structures is the most important point for improvement. This is needed for an adequate read across in (Q)SAR models. This function is financed by BASF but given to the whole community.

- EFSA: We need an organisational form that we should work together in a more coordinated way. The lessons learned from the generic Markush-Structures should be that all components (MetaPath and the composers) should be improved at the same time and the database / XML Files should be migrated but there is only one LMC with limited resources!
- CropLife Europe: It is time to come to the table with all the stakeholders and to clarify, how to work together for the short term improvements. Thinking about needed resources and how industries can help. Organising a workshop where all stakeholders talk about these improvements.
- EFSA: How should such meetings be organised? Should LMC do this? Today they get requests from different sides. And LMC asks US EPA or EFSA asks US EPA and this uncertainty is the reason of a lot of indirect parallel communication.
- CropLife Europe: The MUG should be the group to discuss these problems.
- US EPA: We need a version control of the publicly available MetaPath. May be, that MUG should meet in such an improvement phase much often.
- LMC: Only one version of MetaPath will be supported in the future. The additional identifiers (EFSA) and the generic Markush-structures (BASF) will lead to one public version. This is in consensus with the customer.
- EFSA: The current decision process is ineffective and non-transparent. It needs a lot of communication. We need a tool to collect all the needs, comments, ideas, the needed funding sources to help us to vote for the prioritisation.
- BfR: We know the CropLife Excel list with bugs and improvement proposals including the priority for industry. A compilation of all ideas was already made by EFSA.
 - At first, we need an agreed prioritisation of the open issues provided by the users.
 - then LMC should classify this list according to the resources required and
 - LMC should indicate if there are dependencies between the different issues.One of the most important dependencies is the synchronous maintenance of five different programs. This technology is an artefact from the past. This resource-eating constellation should be overcome as soon as possible.
- EFSA: This list should be optimised to organise a sufficient meantime. All other resources should be invested into the long-term solution.
- USEPA: The next MUG meeting should not be organised before the 3rd week of January. It is proposed to organise meetings at least quarterly.
- EFSA: It could also be difficult in February, because LMC is busy until the end of March.
- BASF: We have a lot of discussion today on IUCLID and the connection between these systems. We should be involved throughout the whole process.
- EFSA: Yes, the main point what we wanted to get from the BfR analysis was an understanding of how to integrate IUCLID and MetaPath together. EFSA hopes to solve the integration challenges. And we should also look to the integration of the (Q)SAR integration of the toolbox.
- BfR: EFSA should use one element of the OpenSource ecosystem, GitHub to collect and manage the issues for the short-term improvement. It would be helpful and provide an impression of how to work together.

Closing the workshop

BfR had asked how to organise the publishing of the meeting report. Is it ok to include the presentations of the other stakeholders into the workshop report? And is a feedback loop needed?

Discussion:

- CropLife Europe: It would be helpful to have such a workshop report. From our side it is not a problem to include the CropLife Europe presentations in the report. It would make sense to organise a feedback loop.
- BfR: So the draft workshop report will be circulated via the MUG mailing list of US EPA.
The understanding that there is the need for an additional meeting shows that these workshops were helpful. Thanks to everybody for their contributions. Thanks for your readiness to participate.
- US EPA: We are grateful that EFSA has moved the MetaPath project forward by establishing a public database and has funded the current improvements.