Establishment of a prototypic Quantitative Microbial Risk Assessment (QMRA) food and feed safety model repository

Estibaliz Lopez de Abechuco, Octavio Mesa Varona, Lars Valentin, Miguel de Alba Aparicio, Thomas Schüler, Taras Günter, Ahmad Swaid, Tasja Buschhardt, Matthias Filter

German Federal Institute for Risk Assessment (BfR)

Abstract

The transfer and re-use of existing food safety knowledge (ranging from experimental data, published mathematical models to risk assessment software code) is currently a major bottleneck in the area of food safety risk assessment. Such knowledge transfer and knowledge re-use would be possible if curated, community-driven risk assessment knowledge repositories would exist, that provide mathematical models/modules in a machine-readable format.

This project therefore aims at the development of necessary resources (standards, ontology-based controlled vocabularies, software tools and services) facilitating the establishment of open, community-driven, curated knowledge repositories for scientists and risk assessors. The main goal of this project is the provisioning of a proof-of-principle implementation of a food and feed safety model repository on the following basis:

- Development of a generic model exchange format (FSK-ML) to encode food and feed safety models, data and simulation scenarios.
- Development of open technical resources (converter, software libraries, data processing workflows) as well as domain-specific ontologies facilitating the broad adoption of FSK-ML in the area of Quantitative Microbial Risk Assessment (QMRA)
- Development and extension of software tools for de-novo generation, export, import, validation and execution of FSK-ML files.
- Establishment of a prototypic web-based knowledge repository for QMRA with seamless interoperation with EFSA's Knowledge Junction (KJ), which allows access to the collected resources as well as the provision of digital object identifiers (DOI).

© German Federal Institute for Risk Assessment (BfR), 2019

Key words: Quantitative Microbial Risk Assessment (QMRA), model repository, information exchange standard, Knowledge Junction (KJ), FSK-Lab, KNIME

Question number: EFSA-Q-2018-00761
Correspondence: AMU@efsa.europa.eu
Disclaimer: The present document has been produced and adopted by the bodies identified above as author(s). In accordance with Article 36 of Regulation (EC) No 178/2002, this task has been carried out exclusively by the author(s) in the context of a grant agreement between the European Food Safety Authority and the author(s). The present document is published complying with the transparency principle to which the Authority is subject. It cannot be considered as an output adopted by the Authority. The European Food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

Acknowledgements: The authors would like to thank the EFSA staff members: Marios Georgiadis, Didier Verloo and Olaf Mosbach-Schulz for the support provided to this scientific output. The authors also wish to thank Esther Sundermann for her valuable contribution and critical review of the final report.

Suggested citation: Lopez de Abechuco E, Mesa Varona O, Valentin L, de Alba Aparicio M, Schüler T, Günter T, Swaid A, Buschhardt T and Filter M, 2019. Establishment of a prototypic Quantitative Microbial Risk Assessment (QMRA) food and feed safety model repository. EFSA supporting publication 2019:EN-1701. 31 pp. doi:10.2903/sp.efsa.2019.EN-1701

ISSN: 2397-8325

© German Federal Institute for Risk Assessment (BfR), 2019
Summary

In December 2016 EFSA and BfR stepped into a Framework Partnership Agreement (FPA) to work jointly on the development, improvement and dissemination of global food safety tools. As outlined in the Specific Agreement #3 of this FPA the specific objective of this work package (WP2) was the development of a prototypic model repository to support the transfer and re-use of existing Quantitative Microbial Risk Assessment (QMRA) models. The newly generated resources should be designed such that a seamless interoperation with EFSA’s Knowledge Junction (KJ) infrastructure is guaranteed. As part of this cooperation, BfR developed new functionalities for their KNIME-based open source software Food Safety Knowledge Lab (FSK-Lab) and created a prototypic web portal “FSK-Web” that allows users to upload, share, find and execute quantitative microbial risk assessment models online. This collaboration also evaluated and contributed to a new harmonized information exchange format, called FSK-ML. This information exchange format can be used to exchange script-based models in an efficient way. Finally, the project generated several web-services that allow to exchange models between the FSK-Web portal and EFSA’s Knowledge Junction.

Key outcomes:

- Improved open information exchange format (FSK-ML) adaptable to various risk assessment domains
- New open source software modules for the KNIME extension FSK-Lab supporting the adoption of FSK-ML by the scientific and risk assessment community
- A prototypic web-based model repository for the exchange of QMRA models including several example models
- Dedicated web-services to generate, edit and execute FSK-ML compliant models online as well as to interoperate with EFSA’s Knowledge Junction (KJ)
- A proof-of-concept for the proposed strategy including a model repository prototype that could be opened publicly in the near future
Table of contents

Abstract ........................................................................................................................................1
Summary ......................................................................................................................................3
1. Introduction ......................................................................................................................5
  1.1. Background and Terms of Reference as Provided by the Requestor ..............................5
  1.2. Interpretation of the Terms of Reference (if Appropriate) Objectives and Working Plan...5
2. Data and Methodologies ..............................................................................................8
  2.1. Infrastructure Development .........................................................................................8
  2.2. Creation of FSK-ML Formatted Models ......................................................................12
    2.2.1. Identification of Sample Models .............................................................................12
    2.2.2. Conversion of Models into FSK-ML Compliant Files ...............................................12
    2.2.3. Modularization of QMRA Models ...........................................................................13
3. Results ............................................................................................................................13
  3.1. Prototypic QMRA Model Repository Food Safety Knowledge Web (FSK-Web) ............14
    3.1.1. FSK_Model_Repository_Demonstrator Workflow ....................................................15
    3.1.2. Upload_of_Harmonized_Models Workflow ...............................................................17
    3.1.3. Online_Creation_of_Harmonized_Models Workflow ...............................................18
    3.1.4. FSKX_Model_Editor Workflow ..............................................................................18
    3.1.5. Information Exchange between EFSA’s Knowledge Junction and FSK-Web ...........19
      3.1.5.1. UploadToZenodo Workflow ..................................................................................19
      3.1.5.2. CreateModelFromZenodo Workflows ....................................................................20
      3.1.5.3. DownloadFromZenodo Workflow .........................................................................21
    3.1.6. Repository Content ...................................................................................................22
3.2. Dissemination, Documentation, Continued Content Generation .....................................24
  3.2.1. Workshops ...............................................................................................................24
3.3. Curation of an Operational Model Repository ...............................................................24
4. Conclusions .....................................................................................................................25
5. Recommendations .............................................................................................................25
References ..................................................................................................................................26
Abbreviations ..........................................................................................................................28
Appendix A – Writing R code compatible with FSK Lab ........................................................29
1. Introduction

1.1. Background and Terms of Reference as Provided by the Requestor

This grant was awarded by EFSA to: Federal Institute for Risk Assessment (BfR)
Beneficiary: Federal Institute for Risk Assessment (BfR)
Grant title: Establishment of a prototypic QMRA food and feed safety model repository
Grant number: GP/EFSA/AMU/2016/01 Specific Agreement #3 (SA3)

1.2. Interpretation of the Terms of Reference (if Appropriate)

Objectives and Working Plan

The food safety community is generating a variety of scientific resources and knowledge relevant for quantitative risk assessments like scientific publications, experimental data, databases, mathematical models as well as software tools. However, the access to these resources and the exchange of information between them are usually difficult and time consuming. This project aims at the establishment of an open community infrastructural framework that will facilitate efficient compilation and exchange of mathematical models relevant for quantitative risk assessments. The objective of the present project is the development of tools (standards, ontology-based controlled vocabularies, software tools and services) that promote an efficient information exchange, transparency and application of good scientific practice within the food safety modelling domain.

As illustrated in Figure 1, Quantitative Microbial Risk Assessment (QMRA) models are currently generated by modelling experts using different software tools, some of which are publicly available while others are commercial products. Some initiatives aim at the creation of databases and model repositories that facilitate the exchange of risk assessment models between different users. However, these initiatives lack the underpinning of an open, harmonized information exchange format that also can serve as a bridge between different software tools. Further, many existing software projects are not designed as open source projects, rendering collaborative software development very difficult. All these issues had been considered and addressed within this EFSA-BfR collaboration project. The objective of this project is to create open infrastructural resources supporting harmonized model annotation, model sharing and the creation of new model-based simulations. This will provide the basis to set up a comprehensive, web-based infrastructure as another foundation for efficient re-use and exchange of models in the food safety domain in the future. For this purpose, the project has generated a prototypic model repository that allows sharing and curating different types of risk assessment models. The establishment of a curated model repository is also expected to promote the adoption of best-practice recommendations for model annotation by the scientific community.
Working Plan

To reach the project goal, new community resources were generated based on two main concepts: harmonization and standardization. The decision on the detailed steps to be carried out within each of the sub-tasks was taken based on the following prioritization:

- **Priority 1**: Technical requirements to accomplish the implementation of a prototypic QMRA model repository
- **Priority 2**: Contractual requirements
- **Priority 3**: Feature requests identified within the requirement analysis task of this project

A brief description of the nature of the work carried out in each sub-task is summarized in Table 1. In Figure 2, a general overview of how the different tasks were integrated in this work package is presented.

**Figure 1**: Overview of the general objective of the project and its most relevant components
Table 1: Description of the main tasks in WP2 of the EFSA-BfR FPA

| Work package | Task description |
|--------------|------------------|
| **WP2.1**  | **2.1.1 Requirement analysis for the QMRA knowledge repository**  
Organisation of a workshop with invited international experts from EFSA, DTU, ANSES, RIVM, University of Cordoba and BfR  
**2.1.2 Refinement and harmonization of the working plan** |
| **WP2.2**  | **2.2.1 Infrastructure development**  
Data standards → improvement of the Food Safety Knowledge Markup Language (FSK-ML) information exchange standard  
QMRA Ontologies → provisioning of ontologies  
Open source software code → development of new features for Food Safety Knowledge Lab (FSK-Lab)  
Web server → development of the required KNIME-server infrastructure  
**2.2.2 Application of FSK-ML to encode existing QMRA modules and models**  
**2.2.3 Integration of FSK-ML standardized models and modules into the QMRA knowledge repository** |
| **WP2.3**  | **2.3.1 Dissemination, documentation, continued content generation (FSK-ML files)**  
Organization of expert workshops, documentation of the created software resources (manuals) generation and updating of FSK-ML formatted QMRA models/modules.  
**2.3.2 Process of establishing an “editorial board” initiated (EFSA)**  
Establishment of guidelines to assure the quality of QMRA models in the QMRA model repository → creation of the Minimum Information Required to Annotate a Risk Assessment Model (MIRARAM) guideline document  
Initiation of discussions on required technical processes, curation procedures and legal aspects of a productive online QMRA model repository  
**2.3.3 Generation of web-based software services**  
Development of prototypic web-services that allow  
Execution, editing, upload and download of QMRA models online (FSK-Web)  
Interoperability between EFSA`s Knowledge Junction and FSK-Web |
2. Data and Methodologies

2.1. Infrastructure Development

One of the key steps to accomplish the project was the development of the necessary infrastructure, which is required for the prototypic QMRA model repository. This project therefore created new solutions in the following areas: “Data standards”, “Open source software tools” and “Other resources supporting proper model annotation (MA)”: 

Data standards: An open information exchange format called “Food Safety Knowledge Markup Language” (FSK-ML) was extended and validated for this project. The original model metadata schema and the controlled vocabularies were proposed from the Risk Assessment Knowledge Integration Platform (RAKIP) community. FSK-ML is an open standard adaptable to different risk assessment domains in a transparent and reproducible way (Haberbeck et al., 2018). FSK-ML and best-practice annotation guidelines were designed such that they can serve as a bridge between different software tools and model repositories in the future. It facilitates the use of models by researchers from different disciplines and thus extends the original application area of the models. FSK-ML defines a file format that allows storing all relevant meta-information on a model in a compressed zip-folder. These specific FSK-ML compliant files carry the .fskx file extension. As represented in Figure 3 such FSKX-files contain the model and visualization scripts, files with the model metadata and the simulation settings and other relevant files like a Manifest.xml, a README.txt, images generated from model-based simulations or publications, etc. To open any FSKX-file the user just has to un-zip the FSKX-file.
Data standards: An open information exchange format called “Food Safety Knowledge Markup Language” (FSK-ML) was extended and validated for this project. The original model metadata schema and the controlled vocabularies were proposed from the Risk Assessment Knowledge Integration Platform (RAKIP) community. FSK-ML is an open XML-based data standard adaptable to different risk assessment domains in a transparent and reproducible way (Haberbeck et al., 2018). FSK-ML and best-practice annotation guidelines were designed such that they can serve as a bridge between different software tools, data and model repositories. Due to this language, FSK-ML formatted data and models can easily be imported, modified and/or exported. This feature facilitates a harmonized way to exchange information, the use of models by researchers from different disciplines and thus extends the original application area of the model. FSK-ML defines a file format that allows storing all relevant meta-information on a model in a compressed zip-folder. These specific FSK-ML compliant files carry the .fskx file extension. As represented in Figure 3 such FSKX-files contain the model and visualization scripts, files with the model metadata and the simulation settings and other relevant files like a Manifest.xml, a README.txt, images generated from model-based simulations or publications, etc. To open any FSKX-file the user just has to un-zip the FSKX-file.

![Model script](https://via.placeholder.com/150)

![Visualization script](https://via.placeholder.com/150)

![Manifest.xml](https://via.placeholder.com/150)

![Figure 3: FSKX-file format for the exchange of models illustrated for a model implemented in the programming language R.](https://via.placeholder.com/150)

An important aspect of the FSK-ML information exchange standard is the adoption of harmonized controlled vocabularies and ontologies. Ontologies are useful resources that could support formal specification of the terms used to annotate models and the relations among them. Ontologies encode standardized terminologies, contextual information and interrelationships of metadata concepts (Griffiths et al., 2017). They can be used by semantic tools or search engines to filter and find relevant models. This project supported the development of the so called FSK-ML “Metadata Master Table” by adopting as many already existing ontologies and controlled vocabularies as possible.

FSK-ML “Metadata Master Table”:

Metadata concepts from the FSK-ML „Metadata Master Table“ ([https://goo.gl/h7YveZ](https://goo.gl/h7YveZ)) were mapped onto external ontologies from the “Bioportal” repository for biomedical ontologies ([https://biopotal.bioontology.org/](https://bioportal.bioontology.org/), Noy et al., 2009) using the “OntoMaton” plugin (Maguire et al., 2013). This plugin facilitates an ontology search and tagging in Google Sheets. Links to the mapped ontologies are specified in column “K” of the “Metadata Master Table”. Ontologies are frequently revised and mapping has, thus, to be updated continuously.
Controlled vocabularies:

A list of controlled vocabularies was provided for as many metadata concepts within the FSK-ML “Metadata Master Table” as possible (https://goo.gl/U4zp7q). Most controlled vocabularies were adopted from the Standard Sample Description for Food and Feed (SSD) Version 2.0 generated by EFSA (2013). Other lists of vocabularies were adopted from domain-specific software tools, such as PMM-Lab (Predictive Microbial Modelling Lab; https://foodrisklabs.bfr.bund.de/pmm-lab/), ICRA (Interactive online Catalogue on Risk Assessment; http://icra.foodrisk.org/) and FDA-iRisk® (Food and Drug Administration Center for Food Safety and Applied Nutrition et al., 2017; https://irisk.foodrisk.org/). Further already existing ontologies that were used are Bibliographic Ontology Specifications for “Publication type” (D’Arcus and Giason, 2009; http://bibliontology.com/), probOnto for “Parameter distribution” (Swat et al., 2016; http://probonto.org), FOODON for “Population name” (https://foodon.org/), and Creative Commons for “Rights” (https://creativecommons.org/).

Open source software tools: Within the project, an open source desktop software solution has been updated to provide an easy-to-use tool to create, import, export, modify or join FSK-ML compliant files. The name of this desktop software is Food Safety Knowledge Lab (FSK-Lab). It is an extension to the Konstanz Information Miner (KNIME) platform (Berthold et al. 2007). KNIME allows users to assemble data analysis tasks in so-called “workflows” that are composed of connected data processing steps called “nodes”, where each node performs a specific task (data access, data mining, statistics, data visualization or reporting, etc.). FSK-Lab extends KNIME through a collection of tailored nodes that allow to perform relevant model annotation and data processing tasks (de Alba Aparicio et al., 2018). FSK-Lab contains nodes that allow users to create, import, export, edit, join and execute FSK-ML compliant files (Figure 4). Further description of the FSK-Lab nodes and the instructions how to download KNIME and FSK-Lab are provided at the FoodRisk-Labs website (https://foodrisklabs.bfr.bund.de/fsk-lab/). Details on the software development progress can be found at the code repository (https://github.com/SiLeBAT/FSK-Lab).
Figure 4: Overview of FSK-Lab nodes and their functionalities

Other resources supporting proper Model Annotation (MA):

MA is a prerequisite for re-usability of models and correct interpretation of model-based predictions. To support modelers in their tasks to properly annotate their models we created a model annotation template for FSK-ML compliant model metadata provisioning. This template is provided as a community online resource (MA-template: https://goo.gl/7adAnS) and was created based on the metadata concepts within the FSK-ML “Metadata Master Table”. The template has the following features:

1. The MA-template allows the user to provide the model metadata online in a Google Sheet, which supports simultaneous model annotation by multiple users. It is also possible to save the template in each user’s Google account. Alternatively, the MA-template can be exported and saved locally as a Microsoft Office Excel file.
2. Microsoft Office Excel and Google Sheets formats of the MA-template are compliant with FSK-ML and can be imported into FSK-Lab by using the "FSK Creator" node.
3. Each metadata concept cell in the MA-template was equipped with a dialog box comprising a definition comment, which aims to contribute to a better understanding of the metadata concepts.
4. Controlled vocabularies for metadata concepts in the MA-template were integrated as dropdown lists. This feature becomes visible once the user creates a copy of the template online in the user’s own Google space or when the user downloads the template as XLSX-file.

5. Cardinality for each metadata concept was provided to specify whether information is considered as mandatory for the annotation of models and data.

6. A direct link was created between the MA-template and the FSK-ML “Metadata Master Table”. This link guarantees that the MA-template is automatically and instantly synchronized whenever changes in the „Metadata Master Table” or the controlled vocabulary lists are performed in the online mode. Note that once the offline Excel sheet is downloaded, the link is broken and the sheet is not updated anymore.

7. A user guide for the MA-template with the title “Model Annotation Excel Template – User Guide” ([https://t1p.de/ieei](https://t1p.de/ieei)) was prepared.

8. A relationship between parameter unit and parameter unit category controlled vocabulary sheets was established to facilitate model annotation in the “Model Annotation Excel Template”. In this way the user does not have to choose the parameter unit category him/herself but it will be added automatically once the parameter unit is chosen.

9. The FSK-ML “Metadata Master Table” contains also a set of sheets with metadata concepts for specific model classes. There are concepts for the following model categories: (1) predictive microbial models, (2) process models, (3) exposure models, (4) dose-response models, (5) health metric models, (6) consumption models, (7) other empirical models, (8) risk characterization models and (9) QMRA models. Models and also data that do not fall into the mentioned categories can be annotated with the “Generic Metadata Schema” sheet of the “Metadata Master Table”. Other properties of the “Metadata Master Table” were described in detail in Haberbeck et al. (2018).

The revision and validation of definitions, data types, metadata concepts and cardinality values within FSK-ML “Metadata Master Table” are ongoing processes. Changes are implemented by taking into account suggestions of all parties involved (i.e. developers, potential users and experts in risk assessment modelling). The ultimate goal is thereby to improve relevance of model metadata and to ease their incorporation into software tools like FSK-Lab. A full list of already achieved improvements is available in the file “Proposals for Improving Harmonization Resources” ([https://goo.gl/3VZ4hM](https://goo.gl/3VZ4hM)).

2.2. Creation of FSK-ML Formatted Models

2.2.1. Identification of Sample Models

During the project a significant effort has been made to identify well annotated open source QMRA models that could be transformed into FSK-ML compliant files and be used in the proof-of-principle implementation of the QMRA model repository.

For the search of models different sources and criteria were used:

- EFSA’s Knowledge Junction (KJ) repository: from the models available in EFSA’s KJ repository the QMRA model on *Listeria monocytogenes* by the EFSA BIOHAZ Panel, 2018 was selected as it complied to the two inclusion criteria: availability of the model R-script and the potential for model modularization.

- Literature review: searched for QMRA models where the code was provided in R.

- Personal communication with QMRA modelling experts

- Software tools with integrated open model repositories
The methods for conversion of R-based QMRA models and modules thereof into FSKX-files are described in the following section (Section 2.2.2).

### 2.2.2. Conversion of Models into FSK-ML Compliant Files

The identified sample models (see Section 2.2.1 for details) had to be converted into FSK-ML compliant files in order to make them available in the prototypic QMRA model repository.

The following steps were performed to transform each model into the FSKX-format:

1. Extraction of model metadata from the original model R code or the original publication and adding it to a model-specific instance of the “Model Annotation Excel Template”. The gathered information includes dependent and independent model parameters, as well as their default values, distributions, minimum and maximum values (if defined and relevant).
2. Execution of the original model in desktop R software to validate whether or not the generated simulation results were consistent with the description provided in the original publication.
3. Refactoring of the original model R script to be compliant with the FSK-ML format considering the “Guidelines for writing R code compliant to FSK-ML” (Appendix A).
4. Application of the desktop software FSK-Lab to convert the re-factored R script and the corresponding “MA-Template” file into FSKX-files. Both files are used as input for the FSK-Lab “FSK Creator” node that integrates all information into an FSK model port object in FSK-Lab.
5. Execution of the created FSK model in FSK-Lab using the “FSK Runner” node and validation of the obtained simulation output with the results provided in the original publication.
6. Saving successfully executed and validated models as FSKX-file using the “FSKX Writer” node. Generated FSKX files contain then among others the following files:
   - Model script
   - Visualization script (R-script for visualization of simulation results)
   - Filled MA-template spreadsheet
   - R workspace generated by the FSK Runner node

### 2.2.3. Modularization of QMRA Models

Risk Assessments are divided in four well-established stages as defined by the Codex Alimentarius Commission (1999): (i) hazard identification, (ii) hazard characterization, (iii) exposure assessment and (iv) risk characterization (Codex Alimentarius Commission, 1999). Certain classes of models find application in specific stages, such as dose-response models (ii), process models and exposure models (iii) and health metric model and risk characterization model (iv). As proposed by Nauta (2001) and extended by Haberbeck et al. (2018), it is advisable to design full QMRA models from a set of model modules that can be combined and adapted according to the specific risk assessment tasks.

As part of this project it has been explored whether it is possible to extract meaningful model modules (e.g. a dose response model module, a microbial growth / inactivation model module, a risk characterization model module etc.) out of the available fully integrated QMRA models. The following steps were performed for that:

1. Thorough inspection of the original model R code and the original model description.
2. Extraction of R code relevant for the specific module from the original R model.
3. Re-factoring of model parameters and code (where necessary).
4. Re-factoring of the visualization script.
5. Annotating of the new model module.
6. Implementing each individual model module as independent FSKX file (as described in Section 2.2.2).

### Results

#### 3.1. Prototypic QMRA Model Repository Food Safety Knowledge Web (FSK-Web)

The prototypic QMRA model repository Food Safety Knowledge Web (FSK-Web) was developed and deployed on the KNIME-server infrastructure of BfR. For this the WebPortal functionality of the commercial KNIME-server software was used, which is available on the BfR server. The KNIME Server WebPortal allows to deploy KNIME workflows created with the open source desktop KNIME Analytics software as interactive browser-based applications. Using the available KNIME-server infrastructure it was possible to design FSK-Web as a set of KNIME workflows (summarized in Table 2) that provide to the user specific web-services to search, explore, execute, edit, upload or download FSK-ML compliant models. FSK-Web contains further additional KNIME workflows to create FSK-ML compliant models online, delete models from the repository and to exchange information with EFSA’s Knowledge Junction (KJ). A dedicated database was not required as FSK-ML compliant model files could be stored directly in a dedicated folder on the KNIME-server.

#### Table 2: Overview of the main workflows and their corresponding functionalities

| Workflow                        | Functionalities                                                                 |
|---------------------------------|---------------------------------------------------------------------------------|
| FSK_Model_Repository_Demonstrator | - Display the model metadata with searching and filtering options               |
|                                 | - Execution of models online with default or user-specific simulation parameters |
| FSK_Model_Downloader            | - Automatic download of FSK-ML formatted models                                  |
| Upload_of_Harmonized_Models     | - Upload of FSK-ML formatted models                                             |
| Online_Creation_of_Harmonized_Models | - Creation of FSK-ML formatted models online                                     |
| FSKX_Model_Editor               | - Editing of existing models in the repository                                   |
|                                 | - Download of the edited version                                                 |
| UploadToZenodo                  | - Automatic submission of a model present in the FSK-Web repository into EFSA’s KJ |
| DownloadFromZenodo              | - Allows the execution of FSK-ML compliant models that are available from EFSA’s KJ using the model’s DOI. If the proper link is specified within the model’s KJ entry the model execution can be initiated directly from the KJ record. |
CreateModelFromZenodo

| CreateModelFromZenodo | Allows to update an existing KJ model record with all metadata collected during the interactive online FSK-ML model generation process. The workflow finally creates and uploads a valid FSK-ML model file based on the user input and generates a new version of the KJ model’s entry that also contains the links required to execute the model via FSK-Web. |

All FSK-Web KNIME workflows contain a start window with an explanation of the workflow and the instructions how to use it. Moreover, on every intermediate page displayed during execution of any KNIME workflow a help button is available, where further instructions are provided. In the following sections each FSK-Web workflow is explained in more detail.

3.1.1. FSK_Model_Repository_Demonstrator Workflow

This workflow provides the main “entry point” or “main view” of the FSK-Web model repository. Once the execution of this workflow is triggered a specific JavaScript-based view of the available models is generated dynamically in the web-browser (see Figure 5 - a detailed description of the repository content is provided in Section 3.1.6). This view shows a tabular collection of models and includes main model metadata (“software”, “environment” and “hazards”) and sophisticated search and filtering functionalities. The full model metadata can be easily accessed via the “Details” button. Analogously, the FSKX-file can be downloaded via the “Download” button.

This main view also provides direct links to other KNIME workflows that are available via the prototypic model repository FSK-Web (see Table 2 for the full list).
Further, the “FSK_Model_Repository_Demonstrator” workflow provides a functionality to execute a model online. It is possible to run simulations with default or user-specific parameter values (see Figure 6 for the simulation parameter settings view). On the left side of this window the user can choose among the available simulations or create a new one by modifying the input parameters listed on the right side. To support the user in designing a new simulation, further information of every parameter is displayed when clicking on the information button, which is labeled with “i”. Among others, this information includes maximum and minimum values (if these are provided by the model creator).
Figure 6: Screenshot of the simulation parameter settings view of the “Full QMRA model for *L. monocytogenes* per type of RTE food product and packaging” model that has been extracted from the original model published by the EFSA BIOHAZ Panel, 2018 to allow analysis of each food product type individually by the type of packaging: normal or Reduced Oxygen Packaging (ROP).

Once a simulation has been created/selected, the model can be executed. The workflow automatically generates a report showing the main settings and the results. The user can download the report (available in different formats .pdf, .docx, .xlsx and .pptx, Figure 7A) and the FSK-ML formatted models, either the model with the default simulation (original model, Figure 7B) or the one created with the user-defined simulation settings (modified model, Figure 7C).
Figure 7: Screenshot of the automatically generated report view. The report can be downloaded in different file formats and contains the links to the original and modified FSK-ML model files. The embedded figures are generated by the so called “visualization script” that the model creator provides together with the model code. All generated numerical simulation results are included into the downloadable FSK-ML files (inside the saved R workspace).

3.1.2. Upload_of_Harmonized_Models Workflow
With this workflow a new FSK-ML compliant model file can be added to the FSK-Web model repository. In a first step the workflow reads in the model metadata from the provided file to retrieve the model information for the main view of the FSK-Web model repository. Then, the model is executed to verify that all required operations (read, run and write out modified FSKX-files) run on the KNIME-server. At the end of these sanity checks, the user can review and update the extracted model metadata and confirm that this model should be added to the FSK-Web model repository.

3.1.3. Online_Creation_of_Harmonized_Models Workflow

FSK-ML compliant model files can be created online through this workflow that provides a web-based graphical user interface (GUI) (Figure 8) so that a user can provide all model metadata and the model code online and create an FSK-ML compliant model file. According to the FSK-ML specifications the requested model metadata cover the four main categories (“General Information”, “Scope”, “Data Background” and “Model Math”) and can be entered via different tabs in the browser window. Additional tabs are available to provide the model script, the visualization script, a README text and supplementary files. A specific feature of FSK-ML is that it allows to differentiate between a model “Author”, who among others generated the model code, and “Creators”, who create the FSKX-file for that model.

Figure 8: GUI to provide model metadata, model and visualization scripts as well as additional resources online.

In Appendix A, detailed instructions are given on how to write the R-code in a way that is compatible with FSK-Lab. In this Appendix there is also a guideline available that describes how to add correctly the model parameters to this GUI.

3.1.4. FSKX_Model_Editor Workflow
The “FSKX_Model_Editor” workflow allows the user to edit an existing FSKX-file in order to update existing metadata and/or change/create new simulation settings and download the new version of the FSKX file afterwards.

### 3.1.5. Information Exchange between EFSA’s Knowledge Junction and FSK-Web

Interoperability between EFSA’s Knowledge Junction and the envisaged QMRA model repository has been an important requirement in this project. With the developed solutions interoperability could be achieved for the three most relevant use cases:

- Automatic submission of models into EFSA’s KJ from FSK-Web
- Integration of model annotation services into EFSA’s KJ publishing process.
- Direct execution of FSK-ML compliant models from within EFSA’s KJ via a dedicated link to a FSK-Web workflow

All developed web-services were tested against the KJ test environment ([http://sandbox.zenodo.org/](http://sandbox.zenodo.org/)) provided by Zenodo and are detailed below.

#### 3.1.5.1. UploadToZenodo Workflow

The “UploadToZenodo” workflow allows users to upload a FSK-ML compliant model from FSK-Web directly to the EFSA’s KJ. The user just needs to select the model in FSK-Web and then this workflow loads the metadata from the model file that are required in EFSA’s KJ (Author, Title, Model Description, Version, Language and Additional Notes) into a new window. If necessary, these metadata can be edited by the user before registering the selected model into EFSA’s KJ. After successful execution of the workflow, the URL of the new KJ entry is provided to the user. The new entry is submitted to Zenodo as candidate for EFSA’s KJ and undergoes a curation process by EFSA before it might be included in the KJ. As shown in Figure 9 the workflow would create a new KJ entry containing:

- Relevant metadata from the FSKX-file (Figure 9A) and a preview of model’s simulation results (Figure 9B)
- An automatically generated DOI (Figure 9C)
- Keywords “FSK-Lab” and “fskx” for easier searching (Figure 9D)
- A link that allows direct execution of the record’s FSKX-file in FSK-Web (Figure 9E)
- The FSKX-file and an image of the model’s simulation results (Figures 9F and 9G, respectively)
Figure 9: Detailed view of the new candidate entry generated in Zenodo from a model in the FSK-Web repository via the UploadToZenodo workflow.

3.1.5.2. CreateModelFromZenodo Workflows

The service provided by the “CreateModelFromZenodo” workflow allows to create a new FSK-ML compliant model using information from an existing entry in EFSA’s KJ. In addition the newly generated FSK-ML compliant model file will be submitted back to EFSA’s KJ as an update to the original KJ entry. Conveniently, this workflow exploits existing information in the original KJ entry to reduce the workload.
for the user. For example, metadata and available R code are automatically integrated into the correct FSK-ML fields. In addition to the existing data, the workflow allows the user to provide further information, e.g., input and output parameters, which are mandatory information. After finalization of the model annotation the original KJ entry will be updated, i.e. it will be re-versioned. Once the workflow has completely executed, both the original and the FSK-ML compliant model are available via the KJ. The new version now contains the updated metadata including the tags ("FSK-Lab" and "fskx") and an URL under “related identifiers” that allow the model to be executed again in FSK-Web. Note that a new DOI from Zenodo is assigned to the updated version of the model. All versions of the model can be tracked as it is illustrated in Figure 10.

Figure 10: Screenshot of a testing example of a KJ item created in Zenodo test environment (Sandbox) that contains a new version created with the CreateModelFromZenodo workflow. The blue box highlights the updated Versions section.

3.1.5.3. DownloadFromZenodo Workflow

This workflow allows to execute FSK-ML compliant models that are available in EFSA’s KJ repository. The execution of the workflow can be triggered either from FSK-Web using the corresponding DOI of an EFSA’s KJ model or through a specific link that can be provided as a “related identifier: supplementary material” in the EFSA’s KJ model entries. To search for FSK-ML compliant models in EFSA’s KJ the users can enter the keywords “FSK-Lab” and “fskx” into the search box. All models uploaded to EFSA’s KJ by the two previously described workflows, namely “UploadToZenodo” and “CreateModelFromZenodo”, contain such proper links. This workflow allows the user to run the model with the default simulation.
settings or create a new simulation by changing the values of the input parameters. Once the model-based simulation has been executed, a new FSKX-file with the new simulation settings and simulation results is created and can be downloaded for further use.

3.1.6. Repository Content

In total, 12 FSK-ML compliant model files were created within this project. The creation of the models provides a proof-of-concept for the developed framework, to test and improve the information exchange format FSK-ML and to test the open source software tool FSK-Lab. All 12 FSK-ML files were generated from existing QMRA models available from different sources as described in Section 2.2.1. As part of this research it was a specific objective to extract useful model modules, e.g. a dose-response module, from available comprehensive QMRA models. Specifically a thorough work was carried out to modularize the *Listeria monocytogenes* generic Quantitative Microbiological Risk Assessment (gQMRA) model (EFSA BIOHAZ Panel, 2018). This model evaluates the number of annual cases of listeriosis in 14 different age and gender groups due to the consumption of different Ready-To-Eat (RTE) food categories (heat-treated meat, smoked and gravad fish as well as soft and semi-soft cheeses) per type of packaging (ROP versus normal). The model is based on an EU-wide baseline survey conducted in 2010 and 2011 to estimate the EU prevalence and contamination levels of *L. monocytogenes* in these RTE food categories. To extract useful model modules from the original gQMRA model, the hazard characterization, exposure assessment and risk characterization elements were identified and converted into fully annotated FSKX-files as described in Section 2.2.3. After performing the necessary reformulation of model parameters, R-code and the creation of new simulation scenarios, the new generated FSKX files facilitated the analysis of RTE foods by single type. Three additional models for Norovirus and other viral pathogens present in water were converted into FSK-ML format and uploaded to the repository.

An overview of all QMRA models and derived model modules is given in Table 3. All generated FSKX model files were successfully integrated into the prototypic FSK-Web model repository and can be executed with default or user-specific simulation settings from within FSK-Web.
Table 3: List of QMRA models and model modules that were made available as FSKX-files. "*" indicates derivative models from the *L. monocytogenes* generic Quantitative Microbiological Risk Assessment (gQMRA) model (EFSA BIOHAZ Panel, 2018).

| Model name | References | Model/module description |
|------------|------------|--------------------------|
| *Listeria monocytogenes* generic Quantitative Microbiological Risk Assessment (gQMRA) | EFSA BIOHAZ Panel, 2018 | *Full QMRA* model for *L. monocytogenes*. The model evaluates the number of annual cases of listeriosis in 14 different age and gender groups due to the consumption of different RTE food categories per type of packaging
|  |  | *Full QMRA* model for *L. monocytogenes* per type of RTE food product and packaging
|  |  | *Dose-Response* model of gQMRA model for *L. monocytogenes* (Hazard Characterization)
|  |  | *Process* model of gQMRA model for *L. monocytogenes* (Exposure Assessment)
|  |  | *Process* model of gQMRA model for *L. monocytogenes* per type of RTE food product and packaging (Exposure Assessment)
|  |  | *Exposure* model of gQMRA model for *L. monocytogenes* (Exposure Assessment)
|  |  | *Exposure* model of gQMRA model for *L. monocytogenes* per type of RTE food product and packaging (Exposure Assessment)
|  |  | *Risk Characterization* model of QMRA model for *L. monocytogenes*
|  |  | *Risk Characterization* model of QMRA model for *L. monocytogenes* per type of RTE food product and packaging type
| *Virus in water - Exposure model* | Haas et al., 2014 | Exposure model for a viral pathogen in water and shellfish (Haas_Example 6.18) Exposure model describing the daily virus exposure from drinking water (Exposure Assessment)
| *Norovirus QRA* | Duret et al., 2017 | Quantitative Risk Assessment of Norovirus transmission in food establishments. The model evaluates the impact of contaminated food, ill/infected employees and their effect on the transmission of Norovirus to foods
|  | Teunis et al., 2008 | Dose-response model for Norovirus (small matrix): Describes the consumers' probability of infection/illness (Hazard Characterization)
3.2. Dissemination, Documentation, Continued Content Generation

3.2.1. Workshops

During the course of the project two workshops have been organized. They were used to evaluate and disseminate the following project results to specialized audiences: (1) FSK-Web portal and (2) the services linking FSK-Web with EFSA's KJ and the FSK-Lab KNIME extension.

1) Title: “Workshop on the EFSA-BfR model repository for Quantitative Microbial Risk Assessment models and its link to the EFSA Knowledge Junction”
   Audience: EFSA staff
   Date: 12.04.2019
   Venue: EFSA

2) Title: “FSK-Web – A PROTOTYPIC QMRA FOOD AND FEED SAFETY MODEL REPOSITORY”
   Audience: Invited experts from EFSA, University of Malta, University of Cordoba, DTU Food, ANSES, IRTA, RIVM, Royal Veterinary College Hatfield, University of Porto, ADRIA Development and BfR.
   Date: 13.05.2019
   Venue: BfR

3.3. Curation of an Operational Model Repository

Within the framework of this project it was also agreed to jointly develop a guideline that could support EFSA and the research community in the curation of the envisaged model repository. Curation is considered as an essential step to assure long-term quality and end-user acceptance for any information repository. A curation process for models aims to classify models according to certain quality criteria. Therefore, such quality criteria must be defined for the envisaged model repository and must then be communicated to the model creators and the users of the repository.

The risk assessment community increasingly supports the idea of establishing guidelines for harmonized annotation of models in jointly agreed exchange formats, as such guidance will improve the quality of model annotations in the long run.

The guideline developed in this project is called MIRARAM. It proposes a minimal set of metadata that needs to be provided by the creator of any FSKX-file in order to be accepted in a model repository. This guidance supports meaningful annotation of mathematical models and mainly focuses on the food safety risk assessment domain.

The MIRARAM guideline is divided into two main parts. Firstly, the general guiding principles are postulated; these principles were adapted from best-practice minimum information guidelines previously established in other scientific disciplines. Secondly, a minimal set of metadata and a specific guidance on technical requirements are provided to enable the fulfillment of the general guiding principles. The minimal set of model metadata was selected from the domain-specific community resource called FSK-ML "Metadata Master Schema" that has been developed and described by food safety risk assessment experts (Haberbeck et al., 2018). Each metadata concept in MIRARAM is defined together with its format and, if applicable, an extendable controlled vocabulary. To avoid future format incompatibility issues, MIRARAM requests that model and model annotations should be provided in files compliant to the COMBINE (Computational Modeling in Biology Network) archive specification defined by the Open Modelling Exchange format (Bergmann, et al., 2014).
We believe that the MIRARAM guidelines could support scientific journal editors/publishers, modelers, risk assessors and software developers in their efforts to make their valuable work re-usable. Further, MIRARAM could also contribute to improve quality and transparency of the model generation process and model-based predictions. It would also help to avoid issues related to intellectual property rights, misinterpretation and potential misuse of data or models. It is planned to publish the MIRARAM guidelines as a scientific paper in the peer-reviewed journal “Microbial Risk Analyses”.

4. Conclusions

This project provided a proof-of-principle implementation of a web-based food safety model repository, called FSK-Web, on the basis of an open information exchange format (FSK-ML), community-driven annotation schema (FSK-ML Metadata Master Schema), open source software solutions (FSK-Lab) and dedicated web-services hosted on BfR’s KNIME-server infrastructure. Specific web-services were developed and deployed that among others facilitate the seamless interoperability of FSK-Web with EFSA’s Knowledge Junction. The evaluation of the developed resources, by external domain experts during two workshops, confirmed the high potential of the generated resources for improving efficient knowledge exchange and quality control in the food safety risk assessment community and beyond. These experts also recommended to increase the number of models in the FSK-Web model repository and to improve usability and performance of the web-based services.

5. Recommendations

Within this project there has been a close interaction with EFSA staff, external risk assessment modeling experts and potential end-users that all influenced and supported the generation of the described project results. Based on the feedback received during the two dissemination workshops it can be recommended to implement the following features in the future:

- Improvements related to programming languages:
  - Support for other scripting languages
  - Support for models created under different R versions

- Implementation of new FSK-Lab nodes:
  - “FSK Joiner” node: to graphically connect two models selecting either input or output parameters from the first model to join them so both are considered as input parameters of the second model. Further it should be possible to create new simulations, run and download the joined models as an FSKX-file. Multiple joins and the joining of models in different scripting languages should be supported.

- FSK-Web features:
  - Improvements on the “Online Generation of FKS-Models” user interface e.g.:
    - Entry of new models showing only relevant metadata based on the model type
    - Multiple authorship
    - Descriptions of every metadata property

  - Improved “Detail view” of the models in the repository:
    - Resizing of metadata fields depending on the content
    - Addition of a new tab with the R-script
- Provision of numerical results in the final report after executing a model
References

Berthold MR, Cebron N, Dill F, Gabriel TR, Kötter T, Meinl T, Ohl P, Sieb C, Thiel K. and Wiswedel B, 2007. KNIME: The Konstanz Information Miner. In Studies in Classification, Data Analysis, and Knowledge Organization (GFKl 2007). Springer.

Codex Alimentarius Commission, 1999. Principles and guidelines for the conduct of microbiological risk assessment. Joint FAO/WHO Food Standards Programme. CAC/GL-Rome

D'Arcus B and Giasson F, 2009. Bibliographic Ontology Specification. Available at: http://bibliontology.com/

de Alba Aparicio M, Buschhardt T, Swaid A, Valentín L, Mesa-Varona O, Günther T, Plaza-Rodriguez C and Filter M., 2018. FSK-Lab – An open source food safety model integration tool. Microbial Risk Analysis, 10, 13-19.

Duret S, Pouillot R, Fanaselle W, Papafragkou E, Liggans G, Williams L and Doren JM, 2017. Quantitative Risk Assessment of Norovirus Transmission in Food Establishments: Evaluating the Impact of Intervention Strategies and Food Employee Behavior on the Risk Associated with Norovirus in Foods. Risk Analysis, 37: 2080-2106.

EFSA (European Food Safety Authority), 2013. Standard Sample Description ver. 2.0. EFSA Journal 2013;11(10):3424, 114 pp. doi:10.2903/j.efsa.2013.3424

EFSA BIOHAZ Panel (EFSA Panel on Biological Hazards), Ricci A., Allende A., Bolton D., Chemaly M., Davies R., Fernández Escámez P. S., Girones R., Herman L., Koutsoumanis K., Narrung B., Robertson L., Ru G., Sanaa M., Simmons M., Skandamis P., Snary E., Speybroeck N., Ter Kuile B., Threlfall J., Wahlíström H., Takkinen J., Wagner M., Arcella D., Da Silva Felício M. T., Georgiadis M., Messens W. & Lindqvist R., 2018. Scientific Opinion on Listeria monocytogenes contamination of ready-to-eat foods and the risk for human health in the EU. EFSA Journal 2018; 16(1):5134,173 pp. doi:10.2903/j.efsa.2018.5134

Food and Drug Administration Center for Food Safety and Applied Nutrition (FDA/CFSAN), Joint Institute for Food Safety and Applied Nutrition (JIFSAN) and Risk Sciences International (RSI). 2017. FDA-iRISK® version 4.0. FDA CFSAN. College Park, Maryland. Available at https://irisk.foodrisk.org/.

Griffiths E, Dooley D, Graham M, Van Domselaar G, Brinkman FSL and Hsiao WWL, 2017. Context Is Everything: Harmonization of Critical Food Microbiology Descriptors and Metadata for Improved Food Safety and Surveillance. Frontiers in Microbiology, 8, 1068.

Haas CN, Rose JB and Gerba CP, 2014. Quantitative Microbial Risk Assessment. Wiley Blackwell.

Haberbeck LU, Rodríguez CP, Desvignes V, Dalgaard P, Sanaa M, Guillier L, Nauta M and Filter M, 2018. Harmonized terms, concepts and metadata for microbiological risk assessment models: the basis for knowledge integration and exchange. Microb. Risk Anal, 10, 3-12.

Nauta, M.J. 2001. A modular process risk model structure for quantitative microbiological risk assessment and its application in an exposure assessment of Bacillus cereus in a REPFED, RIVM report 149106 007.

Noy NF, Shah NH, Whetzel PL, Dai B, Dorof M, Griffith N, Jonquet C, Rubin DL, Storey MA, Chute CG and Musen MA, 2009. BioPortal: ontologies and integrated data resources at the click of a mouse. Nucleic Acids Research. 37 (Web Server issue):W170-W3.
Maguire E, González-Beltrán A, Whetzel PL, Sansone SA and Rocca-Serra P, 2013. OntoMaton: a Bioportal powered ontology widget for Google Spreadsheets. Bioinformatics, 29(4):525-7.

Swat MJ, Grenon P and Wimalaratne S, 2016. ProbOnto: ontology and knowledge base of probability distributions Bioinformatics, 32: 2719-2721

R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Teunis PFM, Moe CL, Liu P, Miller SE, Lindesmith L, Baric RS, Le Pendu J and Calderon RL, 2008. Norwalk virus: How infectious is it? Journal of Medical Virology, 80, 1468-1476.

Weir, MH, Mitchell J, Flynn W and Pope JM, 2017. Development of a microbial dose response visualization and modelling application for QMRA modelers and educators. Environmental modelling & software : with environment data news, 88, 74-83.
Abbreviations

ANSES  L’Agence Nationale de Sécurité Sanitaire de l’alimentation, de l’environnement et du travail - French Agency for Food, Environmental and Occupational Health and Safety
BfR  Bundesinstitut für Risikobewertung - German Federal Institute for Risk Assessment
COMBINE  Computational Modeling in Biology Network archive specification
DTU  Denmark Technical University
EFSA  European Food Safety Authority
FSK  Food Safety Knowledge
FSK-ML  Food Safety Knowledge Markup Language
FSKX  Food Safety Knowledge eXchange file
KJ  Knowledge Junction
KNIME  Konstanz Information Miner
MA  Model Annotation
MIRARAM  Minimum Information Required to Annotate a food safety Risk Assessment Model
PMM-Lab  Predictive Microbial Modelling Lab
QMRA  Quantitative Microbial Risk Assessment
RAKIP  Risk Assessment Modelling and Knowledge Integration Platform
RIVM  Rijksinstituut voor Volksgezondheid en Milieu - National Institute for Public Health and the Environment (Nederland)
ROP  Reduced Oxygen Packaging
RTE  Ready-To-Eat
SSD  Standard Sample Description for food and feed
Appendix A – Writing R code compatible with FSK Lab

Functions not allowed

- `source()` -> All code has to be compiled in one script file.
- `rm(list=ls())`
- `setwd()`, `read()` and `write()`
- No plotting functions in model script

Visualization script

- Only one plot is allowed
- To combine plots → use e.g. `mfrow=c(3,3)`
- Do not use `view()` command
- To convert table/data frames into pictures one can use library(gridExtra) and use function `grid.table(dataframe)`

Importing data

- Currently only `.csv`, `.txt` and `.RData`-files are supported
- Images in .jpg, .bmp, .png and .tiff format can also be uploaded

Parameter annotation

- A “data type” and “default value” has to be provided for each input/constant parameter
- Vectors/matrices/strings can also be defined as input parameters
- Parameters can also be defined as function when parameter “data type” is string
- File input names must be between quotes in “Parameter value” field.
- File input names are case sensitive (e.g. “workspace.Rdata” vs. "Workspace.RData")
- You must use ‘.’ and not ‘,’ for decimal numbers
- Parameter names and IDs should not contain a dot
- Parameter defined as input or constant must have a value.
- Parameters referenced in the default values of other parameters must be defined before being referenced.