Journal Pre-proofs

Minimum Information Required to Annotate food safety Risk Assessment Models (MIRARAM)

Matthias Filter, Esther M. Sundermann, Octavio Mesa-Varona, Tasja Buschhardt, Estibaliz Lopez de Abechuco, Marios Georgiadis

PII: S0963-9969(20)30977-7
DOI: https://doi.org/10.1016/j.foodres.2020.109952
Reference: FRIN 109952

To appear in: *Food Research International*

Received Date: 2 April 2020
Revised Date: 20 November 2020
Accepted Date: 27 November 2020

Please cite this article as: Filter, M., Sundermann, E.M., Mesa-Varona, O., Buschhardt, T., Lopez de Abechuco, E., Georgiadis, M., Minimum Information Required to Annotate food safety Risk Assessment Models (MIRARAM), *Food Research International* (2020), doi: https://doi.org/10.1016/j.foodres.2020.109952

Title

Minimum Information Required to Annotate food safety Risk Assessment Models (MIRARAM)

Authorship

Matthias Filter [a] *, Esther M. Sundermann [a], Octavio Mesa-Varona [a], Tasja Buschhardt [a], Estibaliz Lopez de Abechuco [a] and Marios Georgiadis [b]

Author names and affiliations

Matthias Filter [a, *] (matthias.filter@bfr.bund.de)

Esther M. Sundermann [a] (esther-maria.sundermann@bfr.bund.de)

Octavio Mesa Varona [a] (octavio.mesa-varona@bfr.bund.de)

Tasja Buschhardt [a] (tasja.buschhardt@bfr.bund.de)

Estibaliz Lopez de Abechuco Garrido [a] (Estibaliz.Lopez-de-Abechuco-Garrido@bfr.bund.de)

Marios Georgiadis [b,c] (marios.georgiadis@efsa.europa.eu)

[a] German Federal Institute for Risk Assessment (BfR), Max-Dohrn-Straße 8–10, 10589 Berlin, Germany

[b] European Food Safety Authority (EFSA), Via Carlo Magno 1A, 43126 Parma, Italy

Corresponding author

Matthias Filter (matthias.filter@bfr.bund.de)

[c] Disclaimer

The author Marios Georgiadis is employed with the European Food Safety Authority (EFSA) in the Assessment and Methodological Support Unit, that provides scientific support to all Units and Panels of EFSA. However, the present article is published under the sole responsibility of the authors and may not be considered as an EFSA scientific output. The positions and opinions presented in this article are those of the authors alone and do not necessarily represent the views/any official position or scientific works of EFSA. To know about the views or scientific outputs of EFSA, please consult its website under http://www.efsa.europa.eu.

Declarations of interest:

none

Additional information

All figures should be printed in colour.

Abstract

In the last decades, mathematical models and model-based simulations became important elements not only in the area of risk assessment concerning microbiological and chemical hazards but also in modelling biological phenomena in general. Unfortunately, many of the developed models are published in non-standardized ways, which hinders efficient exchange, re-use and continuous improvement of models within the risk assessment domain. The establishment of guidelines for model annotation is an important pre-condition to overcome these obstacles. Additionally, implementation of annotation guidelines can improve transparency, quality control and even aid the clarification of intellectual property rights. Here, we address the question of "What is the minimum set of metadata that should be provided for a model in the risk assessment domain?". The proposed guideline focuses on food safety risk assessment models and is called "Minimum Information Required to Annotate food safety Risk Assessment Models (MIRARAM)".

MIRARAM supports the model creator during the model documentation step and could also be used as a checklist by scientific journal editors or database curators. Software developers could take up MIRARAM and develop easy-to-use software tools or new features in existing programs that can help model creators to provide proposed model annotations in harmonized file formats. Based on experiences from similar guidelines in related scientific disciplines (like systems biology), it is expected that MIRARAM could contribute to the promotion of application and re-use of models as well as to implementing more standardized quality control in the food safety modelling domain.

Keywords:

Modelling

Quantitative Microbial Risk Assessment (QMRA) Model Annotation checklist

Information exchange format

1. Introduction

According to Codex Alimentarius, risk analysis is a process that comprises three separate elements: (1) risk assessment, (2) risk management and (3) risk communication (FAO/WHO, 2003). Risk assessment determines and identifies hazards and evaluates the risks on human (as well as animal or plant) health via relevant routes of exposure. Risk management establishes the appropriate control measures to monitor and minimize the impact of the risk on the population. Risk communication determines the best way to communicate this information to the affected population (FAO, 1998).

Risk assessments are usually composed of four logically connected sections: (1) Hazard Identification, (2) Hazard Characterization, (3) Exposure Assessment and (4) Risk Characterization. Within Exposure Assessment of hazards in food, mathematical models are often used to describe the fate of the hazard along the relevant exposure routes considering, among other factors, the effect of food process conditions. Modelling and simulation techniques are also frequently applied during Hazard Characterization and Risk Characterization. The complete risk assessments can therefore be composed of several linked "model modules" that estimate diverse characteristics. For example, the prevalence and concentration of the hazard under certain process conditions, the amount of food consumed by distinct population groups, the hazard-specific dose-response relationship (linking the exposure dose to the adverse outcome), the risk linked to a hazard for a certain population group.

One challenge in risk assessment modelling is the lack of harmonized information exchange formats. Guidelines for a harmonized information exchange would allow risk assessors to more easily re-use, adapt and combine existing models created by other risk assessors. Presently, model descriptions are often available as plain-text descriptions without a link to the model implementation. Therefore, in order to use an already published model, often a re-implementation becomes necessary. This old-fashioned way of information exchange is error-prone as errors might not only occur during publishing but also during model re-implementation. For example, consider the names of model parameters; these names are not standardized and therefore the unambiguous identification of model parameters is solely possible from available model documentation/annotation. In short, a successful model re-implementation is at least time-consuming, error-prone and when important information is missing, impossible.

The development of guidelines for model annotation in the risk assessment domain would greatly support efficient model application and re-use. The key element of the guidelines is the definition of a minimum set of metadata that needs to be provided by the model creators. The main purpose of such a minimum metadata set is to ensure that all information critical for model application and re-use is correctly made available. Also, such a minimum information standard can serve to provide quality control criteria for publications on models or model databases.

Minimum information standards are sets of guidelines and formats that are established to support data re-use and data consistency in dedicated scientific communities (Chervitz et al., 2011). They emerged from the need of scientific journals to standardize data publication and to support information exchange between researchers (Chervitz et al., 2011). MIAME (Minimum Information About a Microarray Experiment) (Brazma et al., 2001) became the first initiative that was adopted as a pre-condition for publication in a number of journals, e.g. Nature, Cell and The Lancet (Chervitz et al., 2011). Today, there are several initiatives that deal with the topic of minimum information standards such as MIAPE (Minimum Information About a Proteomics Experiment and all its variants) (Taylor et al., 2007), STRENDA (Standards for Reporting Enzyme Data) (Tipton et al., 2014) or MIRIAM (Minimum Information Requested in the Annotation of biochemical Models) (Le Novère et al., 2005).Adoption of such a standard potentially leads to extensive harmonization of resources in the respective field, as one can see from the MIRIAM guideline that serves as a foundation of the

BioModels Database—a repository of computational models relevant in the context of biological processes (https://www.ebi.ac.uk/biomodels/) (Chelliah et al., 2013; Li et al., 2010). Here, we propose minimum information guidelines for the annotation of models from the area of food safety risk assessment and call them MIRARAM (Minimum Information Required to Annotate a food safety Risk Assessment Model).

2. Methods

In the context of MIRARAM the term *model* is used for the abstract concept of a description of a system through mathematical concepts or software code where the main purpose of such a *model* is to predict one or more properties of that system (therefore also referred to as *predictive model*). The underlying mathematical model or software code can be hidden, i.e., for MIRARAM also so-called “black box” models are within the scope of the *model* definition. A *model* usually encompasses resources, here called *model script / code*, that can be executed by a computer (e.g. files with a software code, a software tool or a web service). A *model script / code* evaluates information provided by the user (here referred to as *model input(s)*), and allows the prediction of defined properties (here referred to as *model output(s)*). *Model input(s)*, *output(s)* and general information about the described *model*, like scope of the model, should be described. Within MIRARAM, the information object that provides all these metadata is called *model annotation or model metadata.* Finally, we use the term *model zip-file* for a digital resource that encompasses the *model annotation* as well as the corresponding *model script / code*.

The proposal for the MIRARAM guideline was developed in collaboration with experts from the risk assessment modelling domain. It is structured into three parts: In the first part, the guideline provides the **general guiding principles**. These principles were adapted from best practices established in similar guidelines of other scientific disciplines, specifically those suggested by Brazma et al. (2001) and Le Novère et al. (2005). These principles guided the selection of specific **metadata requirements** in the second part of MIRARAM. This process aimed at identifying the smallest possible set of model metadata that is needed to fulfil the general guiding principles and support the application and re-use of models. The aspects to be annotated as model metadata are based on an existing comprehensive domain-specific metadata collection, called "Food Safety Knowledge (FSK)" Metadata Schema (Haberbeck et al., 2018). Finally, **technical recommendations** are given to support efficient exchange of models in the future, for example the adoption of the OMEX (Open Modelling Exchange) format specification (Bergmann et al., 2014).

3. Results

3.1. MIRARAM general guiding principles

MIRARAM includes the following four general guiding principles that were adapted from existing minimum information standards to ensure the efficient usage of the provided information*:*

1. The information about a model should be sufficient to interpret it as well as detailed enough to enable comparisons to similar models and to re-use it with other data – adapted from (Brazma et al., 2001).

   This principle implies that model creators have to provide sufficient metadata on model inputs, outputs, model design concept and underlying assumptions to enable comparisons to similar models and to re-use the model with other data.

2. Model metadata should be structured such that it enables automatic data analysis and use – adapted from (Brazma et al., 2001).

   This principle clarifies that model metadata must be provided in an interoperability supporting format, i.e., metadata must be machine-readable as well as in a structured and open format . MIRARAM provides a proposal for such a format, however, there is no obligation to adopt this specific one (see Section 3.3 for details). To comply with this principle, any structured, machine-readable data format capable of representing the model metadata could be used.

3. The model annotation should allow a clear correspondence between a given model documentation and the model itself, i.e., the model must be unambiguously linked to a document that provides all information relevant to model application and interpretation; such a model documentation will be called "reference description" (Le Novère et al., 2005).

   This principle sets the foundation for curation processes. As it is impossible to define general quality control criteria for models, there is only the opportunity to request a dedicated "reference description" so that a reviewer or curator can verify, whether given model metadata are correct and described model-based predictions can be reproduced (Le Novère et al., 2005). As an example, the model creator could reference a specific figure in the "reference description" that a model curator or an end user must be able to reproduce by executing the model with the set of values for the simulation parameters provided by the model creator.

4. The model annotation needs to be correct (Le Novère et al., 2005).

   This principle extends the requirements from Principle 3, as it implies that the annotation also holds true for any application scenarios that would be "allowed" according to the model metadata (Le Novère et al., 2005). For example, if a predictive microbial model has been developed based on experimental data in broth, the model creators have to make sure that the information on the range of applicability of the model states that clearly.

3.2. MIRARAM metadata requirements

MIRARAM defines in this section three specific requirements for model metadata:

1. The *model annotation* must provide information on each of the aspects described in Table 1 (minimum model annotation). For each of the aspects, the annotation can be given as plain text, which means any sequence of ASCII character sequences, e.g. any number, letter or combinations thereof, can be used. The MIRARAM *model annotation* should be created by using *controlled vocabularies* wherever possible (see Table 1).

2. The *model creator* should provide, wherever possible, metadata in addition to the minimum model annotation, to facilitate the re-use and interpretation of the *model*. The FSK Metadata Table is a proposal for a detailed list of model metadata (it is available at https://goo.gl/PE4ysP and in Haberbeck et al. (2018)).This list is considered to be relevant for different food safety model classes, e.g. QMRA models, dose-response models or process models. It is recommended to use *controlled vocabularies* also for these additional, optional *model metadata* (see https://goo.gl/wbFoZU and Haberbeck et al. (2018)).

3. The *model annotation* must be made available in an open, structured, text-based file format (see also Section 3.3).

In Table 2, an example annotation compliant to MIRARAM is provided for a published (in Johne et al. (2016)) predictive microbial model.

3.3 Technical recommendations

MIRARAM recommends that the *model zip-file* (i.e., the *model script / code* and the *model annotation*) is provided inside a ZIP container compliant to the OMEX format (Bergmann et al., 2014). The OMEX format defines that within a ZIP container there needs to be a plain-text file called “manifest.xml” listing all files that are inside this archive. The *model annotation,* as required by MIRARAM, should be provided in a file separate from the *model script / code*. In this way the annotation is not “hidden” in the *model script / code*. This annotation file needs to have a structured, text-based file format, preferably JSON (https://en.wikipedia.org/wiki/JSON) or XML (https://en.wikipedia.org/wiki/XML) - potential names are “modelMetaData.json” or “modelMetaData.xml”. The file(s) comprising the *model script / code* can be included into the ZIP container in multiple ways, e.g. as text file with software code, as a link to a website or as an executable program. It is recommended to provide a dedicated *model script / code* file even if the model is accessible via a web-based resource (i.e., a website or an API) only. A “README.txt” file should be provided inside the ZIP container as well. This file should detail how the *model script / code* can be executed and / or which software / tools / libraries (including version identifier) have to be installed within the required software. The “README.txt” should also be used to provide information for reviewers or end users helping them to verify the concordance of the model based predictions with simulation results presented in the “reference description”. The *model annotation* should include the link to the “reference description” (see Section 3.2). .

4. Discussion

The proposed MIRARAM guideline provides for the first time a minimum set of metadata that a model creator (in the area of food safety risk assessment) should provide for a predictive model. This minimum metadata set was selected from an exhaustive list of model metadata with the aim to reduce the annotation workload while ensuring the provision of all essential information for correct application and re-use of a predictive model by a third person.

In our understanding, the provisioning of requested metadata by the model creator will not require significant additional effort as the requested information should be readily available to them and therefore, easy to provide. The implementation of the technical recommendations from Section 3.3 may, however, create some additional work. Software tools that support users in exporting models into MIRARAM compliant formats are already available, e.g. FSK-Lab (de Alba Aparicio et al., 2018), the R package “FSK2R” (Garre et al., 2019) or the Bioinactivation FE web application (https://foodlab-upct.shinyapps.io/bioinactivationFE/) (Garre et al., 2018). Given the positive effects of guidelines on model annotation in other scientific disciplines, it is our hope that the additional effort required by model creators will result in an overall increase in quality and usability of the model and ultimately will translate into its broader application and re-use by the scientific community. Good indicators of the impact of minimum information standards within a given scientific domain are citation metrics of the publications describing the guidelines, like the field-weighted citation Impact (FWCI) provided by

Elsevier, Scopus ® (https://www.scopus.com/). For example, the publication describing MIRIAM (Le Novère et al., 2005) has a FWCI of 5.64; that for MIAME (Brazma et al., 2001) a value of 2.35 and for MIAPE (Taylor et al., 2007) a value of 21.21. The fast adoption of guidelines suggested by these initiatives and the practical relevance of their derived databases, software tools and frameworks, provide evidence for the effectiveness of the “minimum information” approach.

The main objective of MIRARAM is to describe the minimum information needed to re-use risk assessment models and interpret their outputs. Even though this also necessitates the provisioning of information that supports quality controls and checks on reproducibility of model-based results, it is beyond the scope of MIRARAM to address all challenges linked to the use of a model according to all aspects of the FAIR principles (Wilkinson et al., 2016). An aspect that is outside the MIRARAM scope is the support of interoperability of risk assessment models or of software tools used for modelling. MIRARAM should be considered as one building block that, together with other resources, helps to address those challenges. Resources like the FSK exchange (FSKX) format (see https://foodrisklabs.bfr.bund.de/fsk-ml-food-safety-knowledge-markup-language/ for details) complement MIRARAM. The FSKX format improves the efficiency of information exchange by imposing much stricter technical requirements. This format supports the provisioning of model-related resources like the model annotation, the model script / code, the ReadMe.txt file as well as simulation settings, simulation results or documentations. MIRARAM, together with other resources, should facilitate a more extensive community-wide use of mathematical models in the domain of food safety. There are already several models compliant to MIRARAM that are available via dedicated platforms, e.g. the Virtual Research Environment “RAKIP_portal” (https://aginfra.d4science.org/web/rakip_portal/catalogue) or the Food Modelling Journal (https://fmj.pensoft.net/).

The MIRARAM guideline could be useful for model creators as well as for editors of scientific journals or databases, as it can serve as a checklist for annotations on predictive models. In addition, the guideline gives recommendations on how to provide models in an electronic, machine-readable format (see Section 3.3); specifically we propose to provide a model and all relevant metadata in one “*model zip-file*” (a ZIP container) as also proposed in (Bergmann et al., 2014). MIRARAM could also help to provide clear licence statements and avoid misinterpretation and potential misuse of models (van Panhuis et al., 2014) in its specific area of application.

Acknowledgements

This work was supported by EFSA-BfR Framework Partnership Agreement (FPA) GP/EFSA/AMU/2016/01, Specific Agreement Number #3 (SA3) “Establishment of a prototypic QMRA food and feed safety model repository - WP2” and received funding from the European Union’s Horizon 2020 research and innovation program under the AGINFRA PLUS project (Grant agreement No. 731001).

Abbreviations

DOI - Digital Object Identifier

FSK - Food Safety Knowledge

FMJ - Food Modelling Journal

ID - IDentifier

OMEX - Open Modelling Exchange format

QMRA - Quantitative Microbial Risk Assessment

RAKIP - Risk Assessment Modelling and Knowledge Integration Platforms

REST - REpresentational State Transfer

UID - Unique IDentifier

URI - Uniform Resource Identifier

WPS - Wi-Fi protected Setup

Glossary (source(s) given in parenthesis)

*Controlled vocabulary (CV): "*A set of subject terms, and rules for their use in assigning terms to materials for indexing and retrieval." (http://www.cs.cornell.edu/wya/DigLib/text/Glossary.html).

Field-Weighted Citation Impact (FWCI): The FWCI is a citation metric that allows comparison of the mean article impact across different research areas. A FWCI above 1.00 indicates that the corresponding article is cited more often than the average in its research area. The calculation of the FWCI takes document type, year of publication and the disciplines into account. (https://www.adelaide.edu.au/library/system/files/media/documents/2019-04/ROPE%20Session%20Handout%20Research%20Metrics.pdf, https://doi.org/10.1016/j.joi.2019.03.012)

Metadata: The term *metadata* refers to, "data about data"; in the context of MIRARAM it means information that defines and describes the characteristics of a *model script / code*, used to improve understanding of the model and model-related processes (adapted from https://dictionary.casrai.org/Metadata).

Model: The term *model* is used to refer to the abstract concept of a description of a system through mathematical concepts or software code where the main purpose of such a *model* is to allow the prediction of one or more properties of that system (therefore also referred to as *predictive model*).

Model annotation: The term *model annotation* refers to an information object that provides metadata on *model script / code*. According to MIRARAM *model input(s)*, *output(s)* and general information about the *model*, like the scope, should be provided by the *model annotation.*

Model author refers to one or more person(s) who generated the *model script / code*.

MModel creator refers to one or more person(s) responsible for creating the *model zip-file* in the present form.

Model zip-file: The term *model zip-file* is used to describe a digital resource (i.e. a file) that encompasses at least the *model annotation* and the corresponding *model script / code*.

Model script / code: The term *model script / code* refers to all resources of a *model* that are executable by a computer (e.g. software code, a software tool or a web service) and in this way allow the calculation of the prediction for one or more properties of the modelled system.

Predictive model: The term *predictive model* is used as a synonym for the MIRARAM interpretation of the term *model.*

Reference description: The term *reference description* refers to an unambiguously identifiable information object, e.g. a publication or a file, which provides the most relevant information on the referring model. A *reference description* should provide enough information to allow for some sort of quality control and model validation. For example, the *reference description* could contain a figure or a set of numerical outcomes that should be reproducible when executing the *model script / code* with given input values.

REFERENCES

Bergmann, F.T., Adams, R., Moodie, S., Cooper, J., Glont, M., Golebiewski, M., Hucka, M., Laibe, C., Miller, A.K., Nickerson, D.P., Olivier, B.G., Rodriguez, N., Sauro, H.M., Scharm, M., Soiland-Reyes, S., Waltemath, D., Yvon, F. and Le Novere, N., 2014a. COMBINE archive and OMEX format: one file to share all information to reproduce a modeling project. *BMC Bioinformatics*. 15, 369. 10.1186/s12859-014-0369-z

Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., Aach, J., Ansorge, W., Ball, C.A., Causton, H.C., Gaasterland, T., Glenisson, P., Holstege, F.C.P., Kim, I.F., Markowitz, V., Matese, J.C., Parkinson, H., Robinson, A., Sarkans, U., Schulze-Kremer, S., Stewart, J., Taylor, R., Vilo, J. and Vingron, M., 2001. Minimum information about a microarray experiment (MIAME)—toward standards for microarray data. *Nature Genetics*. 29, 365. 10.1038/ng1201-365

Chelliah, V., Laibe, C. and Le Novère, N., 2013. BioModels Database: A Repository of Mathematical Models of Biological Processes, in: Schneider, M.V. (Ed.), *In Silico Systems Biology*. Humana Press, Totowa, NJ, pp. 189-199.

Chervitz, S.A., Deutsch, E.W., Field, D., Parkinson, H., Quackenbush, J., Rocca-Serra, P., Sansone, S.-A., Stoeckert, C.J., Taylor, C.F., Taylor, R. and Ball, C.A., 2011. Data Standards for Omics Data: The Basis of Data Sharing and Reuse. *Methods in molecular biology* (Clifton, N.J.). 719, 31-69.

de Alba Aparicio, M., Buschhardt, T., Swaid, A., Valentin, 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. 10.1016/j.mran.2018.09.001

FAO, 1998. Food Quality and Safety Systems - A Training Manual on Food Hygiene and the Hazard Analysis and Critical Control Point (HACCP) System. Annex 2-The application of risk analysis to food safety control programmes. Publishing Management Group, FAO Information Division.

FAO/WHO, 2003. CODEX ALIMENTARIUS COMMISSION. Appendix IV. Working Principles for Risk Analysis for Application in the Framework of the Codex Alimentarius.

Garre, A., Clemente-Carazo, M., Fernández, P.S., Lindqvist, R. and Egea, J.A., 2018. Bioinactivation FE: A free web application for modelling isothermal and dynamic microbial inactivation. Food Research International. 112, 353-360.

Garre, A., Peñalver-Soto, J.-L., de Alba Aparicio, M., Egea, J.A., Fernandez, P.S. and Filter, M., 2019. FSK2R: A NEW R LIBRARY TO SUPPORT FOOD SAFETY KNOWLEDGE MARKUP LANGUAGE (FSK-ML), 11th International Conference on Predictive Modelling in Food. Braganza, Portugal.

Haberbeck, L.U., Plaza-Rodríguez, C., 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. *Microbial Risk Analysis*. 10, 3-12. 10.1016/j.mran.2018.06.001

Johne, R., Trojnar, E., Filter, M. and Hofmann, J., 2016. Thermal Stability of Hepatitis E Virus as Estimated by a Cell Culture Method. *Appl Environ Microbiol.* 82, 4225-4231.

Le Novère, N., Finney, A., Hucka, M., Bhalla, U.S., Campagne, F., Collado-Vides, J., Crampin, E.J., Halstead, M., Klipp, E., Mendes, P., Nielsen, P., Sauro, H., Shapiro, B., Snoep, J.L., Spence, H.D. and Wanner, B.L., 2005. Minimum information requested in the annotation of biochemical models (MIRIAM). *Nature Biotechnology*. 23, 1509. 10.1038/nbt1156

Li, C., Donizelli, M., Rodriguez, N., Dharuri, H., Endler, L., Chelliah, V., Li, L., He, E., Henry, A., Stefan, M.I., Snoep, J.L., Hucka, M., Le Novère, N. and Laibe, C., 2010. BioModels Database: An enhanced, curated and annotated resource for published quantitative kinetic models. *BMC Systems Biology.* 4, 92.

McMurry, J.A., Juty, N., Blomberg, N., Burdett, T., Conlin, T., Conte, N., Courtot, M., Deck, J., Dumontier, M., Fellows, D.K., Gonzalez-Beltran, A., Gormanns, P., Grethe, J., Hastings, J., Hériché, J.-K., Hermjakob, H., Ison, J.C., Jimenez, R.C., Jupp, S., Kunze, J., Laibe, C., Le Novère, N., Malone, J., Martin, M.J., McEntyre, J.R., Morris, C., Muilu, J., Müller, W., Rocca-Serra, P., Sansone, S.-A., Sariyar, M., Snoep, J.L., Soiland-Reyes, S., Stanford, N.J., Swainston, N., Washington, N., Williams, A.R., Wimalaratne, S.M., Winfree, L.M., Wolstencroft, K., Goble, C., Mungall, C.J., Haendel, M.A. and Parkinson, H., 2017. Identifiers for the 21st century: How to design, provision, and reuse persistent identifiers to maximize utility and impact of life science data. *PLoS biology*. 15, e2001414-e2001414. 10.1371/journal.pbio.2001414

Palisade Corporation, 2006. @RISK. Palisade Corporation, New York [USA].

Python Software Foundation, 2019. Python Language Reference.

R Development Core Team, 2010. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Taylor, C.F., Paton, N.W., Lilley, K.S., Binz, P.-A., Julian Jr, R.K., Jones, A.R., Zhu, W., Apweiler, R., Aebersold, R., Deutsch, E.W., Dunn, M.J., Heck, A.J.R., Leitner, A., Macht, M., Mann, M., Martens, L., Neubert, T.A., Patterson, S.D., Ping, P., Seymour, S.L., Souda, P., Tsugita, A., Vandekerckhove, J., Vondriska, T.M., Whitelegge, J.P., Wilkins, M.R., Xenarios, I., Yates Iii, J.R. and Hermjakob, H., 2007. The minimum information about a proteomics experiment (MIAPE). *Nature Biotechnology.* 25, 887.

Tipton, K.F., Armstrong, R.N., Bakker, B.M., Bairoch, A., Cornish-Bowden, A., Halling, P.J., Hofmeyr, J.-H., Leyh, T.S., Kettner, C., Raushel, F.M., Rohwer, J., Schomburg, D. and Steinbeck, C., 2014. Standards for Reporting Enzyme Data: The STRENDA Consortium: What it aims to do and why it should be helpful. *Perspectives in Science*. 1, 131-137.

van Panhuis, W.G., Paul, P., Emerson, C., Grefenstette, J., Wilder, R., Herbst, A.J., Heymann, D. and Burke, D.S., 2014. A systematic review of barriers to data sharing in public health. *BMC Public Health*. 14, 1144.

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J., Groth, P., Goble, C., Grethe, J.S., Heringa, J., t Hoen, P.A., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J. and Mons, B., 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data*. 3, 160018.

Web references

D4Science. 2020. AGINFRA RAKIP_PORTAL Catalogue. https://aginfra.d4science.org/web/rakip_portal/catalogue (accessed 20 November 2020).

European Food Safety Authority.2020. http://www.efsa.europa.eu (accessed 20 November 2020).

EMBL-EBI, 2020. BIOMODELS. https://www.ebi.ac.uk/biomodels/ (accessed 20 November 2020).

Field-weighted citation Impact (FWCI) provided by Elsevier, Scopus ®. https://service.elsevier.com/app/answers/detail/a_id/14894/ (accessed 16th October 2020).

German Federal Institute for Risk Assessment (BfR), 2019. FoodRisk-Labs. https://foodrisklabs.bfr.bund.de/frl/ (accessed 20 November 2020).

German Federal Institute for Risk Assessment (BfR), 2019. FSK-ML (Food Safety Knowledge Markup Language). https://foodrisklabs.bfr.bund.de/fsk-ml-food-safety-knowledge-markup-language/ (accessed 20 November 2020).

German Federal Institute for Risk Assessment (BfR), 2019. Controlled_Vocabularies_Master_Table_V1.04. https://goo.gl/wbFoZU (accessed 20 November 2020).

German Federal Institute for Risk Assessment (BfR), 2019. Metadata_ Master_Table_V1.04. https://goo.gl/PE4ysP (accessed 20 November 2020).

Pensoft Publishers, 2019. FMJ Food Modelling Journal. https://fmj.pensoft.net/ (accessed 20 November 2020).

Technical University of Cartagena. 2018. Bioinactivation FE. https://foodlab-upct.shinyapps.io/bioinactivationFE/ (accessed 30 March 2020).

Wikipedia – The Free Encyclopedia. https://en.wikipedia.org/wiki/JSON, https://en.wikipedia.org/wiki/XML (accessed 16th October 2020).

Tables

Table 1: Metadata that must be provided in a structured text-based data format to achieve MIRARAM compliance. "*" indicates the availability of controlled vocabularies for this concept at https://goo.gl/wbFoZU (see also Haberbeck et al. (2018)).

Metadata concept	**Definition of the concept**
Model name	A name given to the model
Model ID	A unique identifier (McMurry et al., 2017) allowing a *model zip-file* to be distinguished from any other uniquely identified information. The identifier can be an auto-generated unique identifier (UID), a

	uniform resource identifier (URI), a digital object identifier (DOI) or any other string that guarantees uniqueness.
Model creator	The person(s) that provide(s) the *model zip-file.* Note that the *model creator* might differ from the model author, which is the person who developed the model (generated the *model script / code)*.
Creation date	Creation date/time of the *model zip-file*
License	Rights granted by the model creator on usage, distribution and modification of the *model zip-file*. The model creator is responsible to respect any rights of others, e.g. the rights of model author(s) or journals in case the model code was retrieved from a publication (including, for example all relevant licenses and copyrights) .
Model execution	A descriptor that facilitates future interoperability (preferably from a controlled vocabulary). In the case of models provided as software scripts (e.g. written in R (R Development Core Team, 2010) or Python (Python Software Foundation, 2019)), this metadata field would describe the name and version of the scripting language. In the case of models provided as executable tools, this field would identify required tools and execution environments. In case of web services, this field would identify web protocols (e.g. REST or WPS).
Reference description	URI or pointer to the resource in which the model has been described in detail and that can be used for quality control. If the license allows and all relevant rights (e.g. copyright) are respected it is recommended to include the reference description into the *model zip-file.*
Model scope	General description on the scope, purpose and application limits of the model.
The following metadata should be provided FOR EACH relevant model input and output. It is up to the model creator to decide if a model input or output is considered relevant for the end user and should therefore be annotated.	
Parameter ID	An unambiguous (inside the *model zip-file*) ID given to the parameter (could be auto-generated by the modelling tool).
Parameter classification *	Classification of the parameter (Input, Constant, Output).
Parameter unit	Unit of the parameter.
Parameter data type *	Information on the data format of the parameter.
Parameter default value	Parameter values allowing model execution. It is recommended to set Parameter default values such that model execution leads to results described in the “reference description”, e.g. in figures.

	Please note: this requirement only applies to parameters classified as “Input” or “Constant”!

Table 2: Example of a model annotation compliant to the MIRARAM guideline that describes an inactivation model of *HEV* as published in Johne et al. (2016).

Model name	HEV-Infectivity_CellSuspension_70C-Temp_InactivationModel
Model ID	HEV-Infectivity_CellSuspension_70C-Temp_InactivationModel
Model creator	Matthias Filter
Creation date	January 2016
Licence	CC BY-NC-SA
Model execution	PMM-Lab
Reference description	10.1128/AEM.00951-16
Model scope	Predictive Model on HEV inactivation in cell culture medium at 70°C for up to 90 sec
Model math	$\mathrm{Value} = \mathrm{LOG10N0} - \left(\frac{\mathrm{Time}}{delta}\right)^{\mathrm{p}}$
Parameter ID	Value
Parameter classification	Output
Parameter unit	Log10(ffu/ml)
Parameter data type	Double
Parameter ID	LOG10N0
Parameter classification	Input
Parameter unit	Log10(ffu/ml)
Parameter data type	Double
Parameter default value	3.9
Parameter ID	delta
Parameter classification	Constant
Parameter unit	d
Parameter data type	Double
Parameter default value	0.0002
Parameter ID	p

Parameter classification	Constant
Parameter unit	[]
Parameter data type	Double
Parameter default value	0.79
Parameter ID	Time
Parameter classification	Input
Parameter unit	d
Parameter data type	Double
Parameter default value	0.001

Author contribution

Matthias Filter: Supervision, Project administration, Conceptualization, Methodology, Writing - Original Draft, Writing - Review & Editing

Esther M. Sundermann: Writing - Original Draft, Writing - Review & Editing

Octavio Mesa-Varona: Conceptualization, Methodology, Writing - Original Draft

Tasja Buschhardt: Writing - Original Draft, Writing - Review & Editing

Estibaliz Lopez de Abechuco: Writing - Original Draft, Writing - Review & Editing

Marios Georgiadis: Supervision, Writing - Review & Editing

Declarations of interest

Title

Minimum Information Required to Annotate food safety Risk Assessment Models

(MIRARAM)

Authorship

Matthias Filter [a], Esther M. Sundermann [a], Octavio Mesa-Varona [a], Tasja Buschhardt [a],

Estibaliz Lopez de Abechuco [a] and Marios Georgiadis [b]

Author names and affiliations

Matthias Filter [a, *] (matthias.filter@bfr.bund.de)

Esther M. Sundermann [a] (esther-maria.sundermann@bfr.bund.de)

Octavio Mesa Varona [a] (octavio.mesa-varona@bfr.bund.de)

Tasja Buschhardt [a] (tasja.buschhardt@bfr.bund.de)

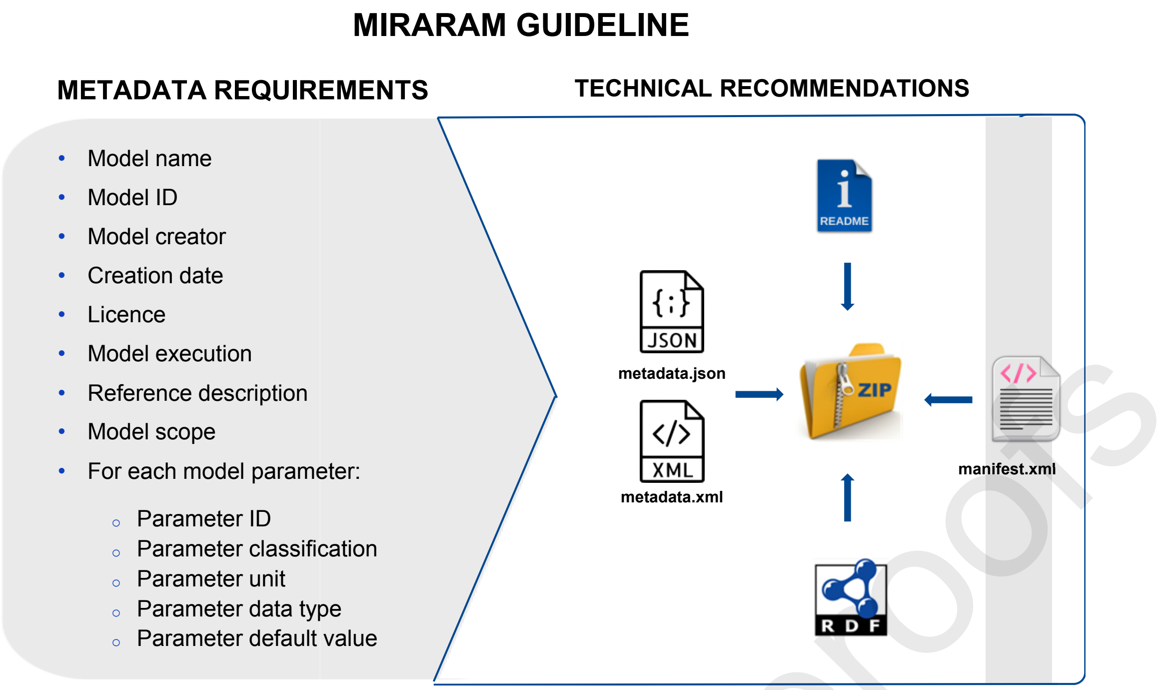
Estibaliz Lopez de Abechuco Garrido [a] (Estibaliz.Lopez-de-Abechuco-Garrido@bfr.bund.de)

Marios Georgiadis [b] (marios.georgiadis@efsa.europa.eu)

[a] Federal Institute for Risk Assessment (BfR), Unit Food Technologies, Supply Chains and Food Defense, Department 4 - Biological Safety, Max-Dohrn-Straße 8–10, 10589 Berlin, Germany

[b] European Food Safety Authority (EFSA) - Risk Assessment and Scientific Assistance Department (RASA), Via Carlo Magno 1A, 43126 Parma, Italy

Declarations of interest: none


MIRARAM GUIDELINE
METADATA REQUIREMENTS
TECHNICAL RECOMMENDATIONS
Model name
Model ID
Model creator
Creation date
Licence
Model execution
Reference description
Model scope
For each model parameter:
Parameter ID
Parameter classification
Parameter unit
Parameter data type
Parameter default value
README
JSON
metadata.json
XML
metadata.xml
ZIP
manifest.xml
R D F

Highlights

- Most food safety and risk assessment models still omit essential metadata
- Guidelines on how to annotate models in a harmonized way are required
- Proposed MIRARAM guideline builds on best practices from related disciplines
- Adoption of MIRARAM will facilitate model re-use and quality control