The development of a decision support system for the infant food chain

Brian Flynn, Rallou Thomopoulos

Global Pediatrics 2 (2022) 100015

Contents lists available at ScienceDirect

Global Pediatrics

journal homepage: www.elsevier.com/locate/gpeds

The development of a decision support system for the infant food chain

Brian Flynn [a,*], Rallou Thomopoulos [b]

[a] Creme Global, The Design Tower, Trinity Technology and Enterprise Campus, Grand Canal Quay, Dublin 2, Ireland
[b] University of Montpellier, INRAE, Institut Agro, IATE, Montpellier, France

ARTICLE INFO

Keywords:
Food safety
Baby food
Chemical contaminants
Foodborne hazards
Risk assessment
Hazard control
Hazard detection
Decision support system

SUMMARY

In a time where the awareness of food safety and quality increases among the general population, it is vital that consumers are enabled to make informed decisions on risks involving the safety of their food. The SAFFI (Safe Food for Infants in EU and China) project aims to build an integrated decision support system (DSS) for the infant food chain that will enable stakeholders at all levels to make informed decisions regarding infant food. The infant food chain was selected due to its strict regulatory requirements, its vulnerabilities as highlighted by different food safety crises, the economic importance of the infant food sector in the EU and China and the focus on this particular food chain by food safety authorities.

The SAFFI project will incorporate data and models from work packages dealing with hazard identification (HI), hazard detection (HD), hazard control (HC) and risk ranking (RR). The models will be integrated into a user-friendly and upgradeable cloud-based decision support system application. A multi-actor cost-benefit analysis of the project will be carried out, enabling the stakeholders in the project to assess the relevance of implementing the project technologies by integrating food safety, regulatory and economic criteria.

The decision support system will be validated on four specific case studies, and tested on end-users, with the aim of extending this approach to other food chains.

1. Introduction

As the awareness of food safety and quality increases among the general population, it is of paramount importance that consumers are enabled to make better informed decisions regarding the risks they take in their everyday life. In the era of big data, in which the number of research papers on food data has grown nearly 300% every five years since 2010,[18] there has never been a better time to harness this data in such a way to manage the issue of food safety. The SAFFI (Safe Food for Infants in EU and China) project aims to build an integrated decision support system (DSS) for the infant food chain that will enable stakeholders at all levels to make informed decisions regarding infant food. The infant food chain was chosen as the focus of the SAFFI project for several reasons: (1) if the DSS tool can successfully integrate the various elements monitoring the high standards of safety that infant food producers must adhere to given the vulnerability of the population, it can then be expanded to other food chains; (2) the number of high-profile incidents in this supply chain indicate the focus that infant food is given and the susceptibility of this chain to food safety incidents; (3) The most recent French total diet study focused specifically on the infant population, demonstrating the importance of this population to authorities. Finally, the infant food industry is of great importance both in the EU (the sixth most valuable product category exported in 2021[9]) and China, where the growth of the sector has been strong.

The main priorities of SAFFI project are: (i) to have a better insight on microbiological and chemical hazards along the infant food chain; (ii) to identify the main known risks and provide (when needed) new tools for their identification, detection, assessment and mitigation by both public health authorities and food industry; (iii) to anticipate unknown risks related to chemical contaminants not detected by current monitoring systems; (iv) to prevent public health crises related to foodborne microorganisms by proposing tools for predictive microbiology and risk management based no longer on hazards but on risks; (v) to further share data, practices, and critical information in real time to ensure overall food safety control.

The activities based on these priorities will culminate in an integrated decision support system. A decision support system can be described as the "*hardware/software that allows a specific decision maker or group of decision makers to deal with a specific set of related problems*".[17] The anticipated decision makers (i.e., the users of the DSS) will range

* Corresponding author.
E-mail address: brian.flynn@cremeglobal.com (B. Flynn).

https://doi.org/10.1016/j.gpeds.2022.100015

Available online 2 May 2022

from the preparer of infant food (who, via the DSS, will be enabled to make safer decisions around food use) to infant food companies and food safety authorities, who will be enabled to make appropriate responses when a new product is developed, or a new hazard is identified or suspected in the infant food chain. As part of work package 5, a multi-actor cost-benefit analysis of the decision support system will be carried out. This cost-benefit analysis will enable the different stakeholders in the project to assess the relevance of implementing the project technologies by integrating food safety, regulatory and economic criteria. By assessing this, SAFFI's cost-benefit analysis will enable the project partners to directly assess how the project outcomes will impact the performance and reliability of food safety control all along the infant food chain.

The decision support system will be developed in the form of a beta software programme, and in adherence to the Technology Readiness Levels detailed by the Horizon 2020 programme,[8] will reach a TRL of 4, signifying the technology developed has been validated in a lab. This beta software will be further developed, upon the completion of the SAFFI project, with the goal of producing a commercially ready DSS tool that can be adapted to other food chains and upgraded with new data in the future.

2. Methods

2.1. Decision support system

2.1.1. Data collection

There are seven work packages within the SAFFI project, the first four of which will produce data that forms the basis of the integrated decision support system (DSS) to be developed. These work packages are individually concerned with unique aspects of food safety, including hazard identification (HI), hazard detection (HD), hazard control (HC) and risk ranking (RR). The types of data being generated and collected for use in the integrated DSS are equally unique, ranging from food processing data from pilot experiments to metagenomics data to cost-benefit analysis data. This data, and the subsequent models which it will be incorporated into for the integrated DSS, will be hosted on the SAFFI Data Foundry platform.[5] SAFFI Data Foundry is a cloud based platform developed by Creme Global, Dublin, Ireland, that facilitates the secure collection of datasets, hosts data collection portals and will ultimately host the integrated DSS developed during the SAFFI project. As there is a vast quantity and variety of data being collected it is crucial to the model development process that the data collected is of a high quality and is suitable for integration into a computational model. The data collection within this project needs to be considered on two levels: (1), the data collection process must be managed by Creme Global to ensure that the templates used for data collection and the databases built must be of a suitable standard to be integrated into a computational model, and an integrated DSS; (2) under the Open Research Data Pilot and EU Horizon 2020 project guidelines,[7] project partners are obliged to follow the FAIR principles in making data Findable, Accessible, Interoperable and Reusable. The steps taken to ensure the data produced in this project will meet the FAIR requirements include the use of accepted protocols in recording metadata and labeling datasets, the storage of raw data in a data repository and the agreement of a data sharing agreement (to be established as the project continues), the publication of results in an OpenAccess journal where possible, and the following of OpenAIRE guidelines for online interoperability of results. As part of the SAFFI project, a Data Management Plan has been developed and will be updated and referred to as the project progresses.

2.1.2. Requirements gathering

In order to build models that satisfy the technical requirements of the respective work packages and the overarching technical objectives of the project, the requirements of the model to be developed must be identified. The purpose of the requirements gathering step is to scope out the requirements of the model- including the use cases of the model, the functional requirements / features, the inputs and outputs, possible constraints and interface specifications- and to allow the identification of key data inputs. The key data inputs, as well as the actual data to be incorporated into the model, include the qualitative and quantitative procedures that will be used to evaluate the content of databases, and the mathematical models, algorithms and decision-based models that will be applied to data. To commence the requirements gathering process, surveys will be circulated among the partners in the respective work packages. These surveys will collect, from the partners, descriptions of the proposed models, essential inputs, desired outputs and algorithms underlying the running of the model. To steer the requirements gathering process, close collaboration is required between the model developer and the partners involved in that particular model.

2.1.3. Model development (I)

Data collection templates and databases developed by the SAFFI partners are profiled to ensure that they are suitable for integration into a computational model. Data profiling involves the analysis of the data collected or generated in order to assess the quality of the data, to clean data or to identify gaps, to create metadata, identify dependencies between datasets or to develop schema.[14]

Following this collection and profiling of data, and the collection of model requirements, initial data modeling is performed. This involves the development of conceptual and logical data models[24] using entity relationship diagrams (ERDs) and data-flow diagrams. Entity relationship diagrams are a type of flowchart used to visualize how "entities", in this case datasets and databases, relate to one another in a system, while data-flow diagrams represent the flow of data through a process. Both diagrams will be used in the model development process to position the collected datasets and databases into a database schema that mathematical and decision based models, and software, can interpret and rely on.

The diagrams designed for each model can be used to optimize the model schema for the most efficient performance in terms of speed to run the model and computational load, essentially optimizing the model to allow the least number of calculations to be required for the desired output.

Following the development of conceptual models for each model, and the subsequent optimisation, each proposed model can be evaluated to identify the concept(s) most suitable to progress further in the model development process. Within the SAFFI project, this suitability to progress criteria will involve collaboration between the respective experts in work packages 1 to 4 and the software development experts in work package 5. Aspects to be considered will be the functional requirements of the model, the realistic capabilities within the timeframe of the project, and the limitations of the data that is available to the project.

2.1.4. Model development (II)

When the models to be developed have been selected, more in-depth data profiling and mapping must take place, including a greater inter-dataset analysis and organization of key variables within the datasets. Again, interdependencies will be identified between datasets and even between models, and a logical data model must be designed to visualize the flow of data through the envisaged model. Based on the finalized design of the logical data model, a physical data model will be built. The datasets that had been generated and profiled previously will be adapted and organized to fit the model, this will involve having separate tables that serve as inputs and variables to the underlying algorithms of the model. Data used by the model will be accessible to the user, and editable or updateable as the user requires.

2.1.5. Testing and deployment

The beta software tool will be tested via scenarios with known results and outputs, in collaboration with the other partners. This testing will involve running the model using inputs from an experiment that has

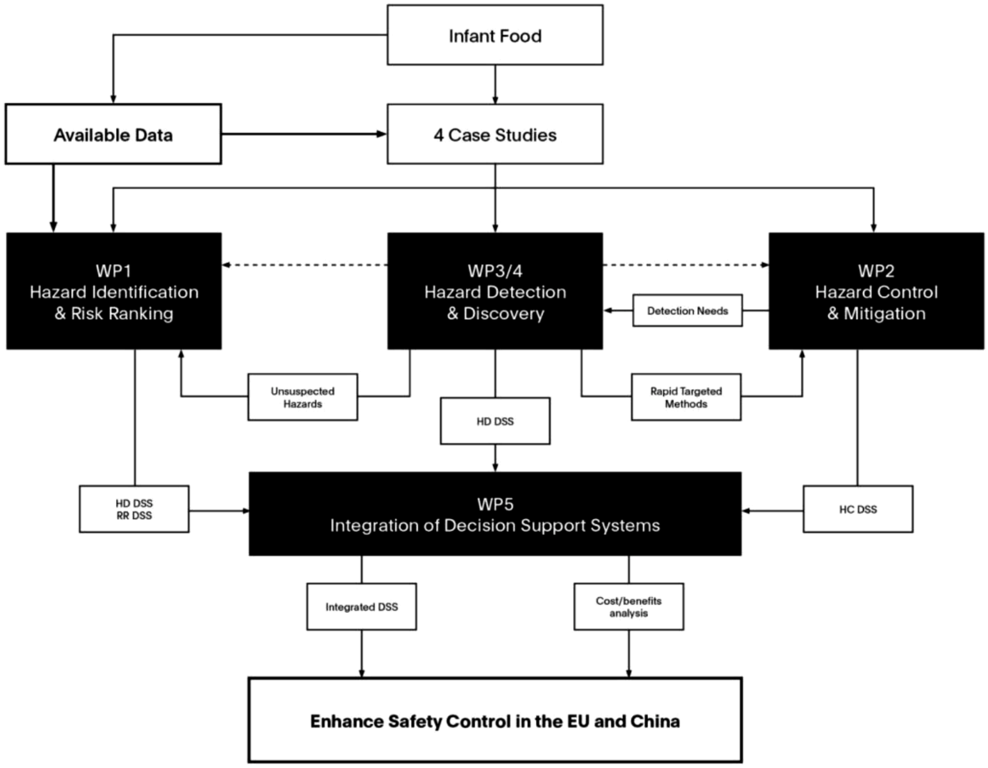

Fig. 1. SAFFI project structure and work package interdependencies.

physically been carried out and run, and the outputs that the model generates will be compared to the physical results so as to validate its effectiveness.

The model will be hosted on Data Foundry, a cloud-based software tool that enables the use of complex data science products behind user-friendly interfaces through a web browser. Creme Global's proprietary technology accommodates the processing loads required by the complex data models that will be developed through the SAFFI project. Functionalities of Data Foundry includes account management, data editors, file management systems, data up-loaders and modeling engines. The platform uses Amazon Web Services allowing for a single point of entry that is highly secure. Reproducibility is achieved by keeping copies of the essential data for re-creating scenarios. Collaboration between parties is enabled by the cloud-based nature of the platform, and the

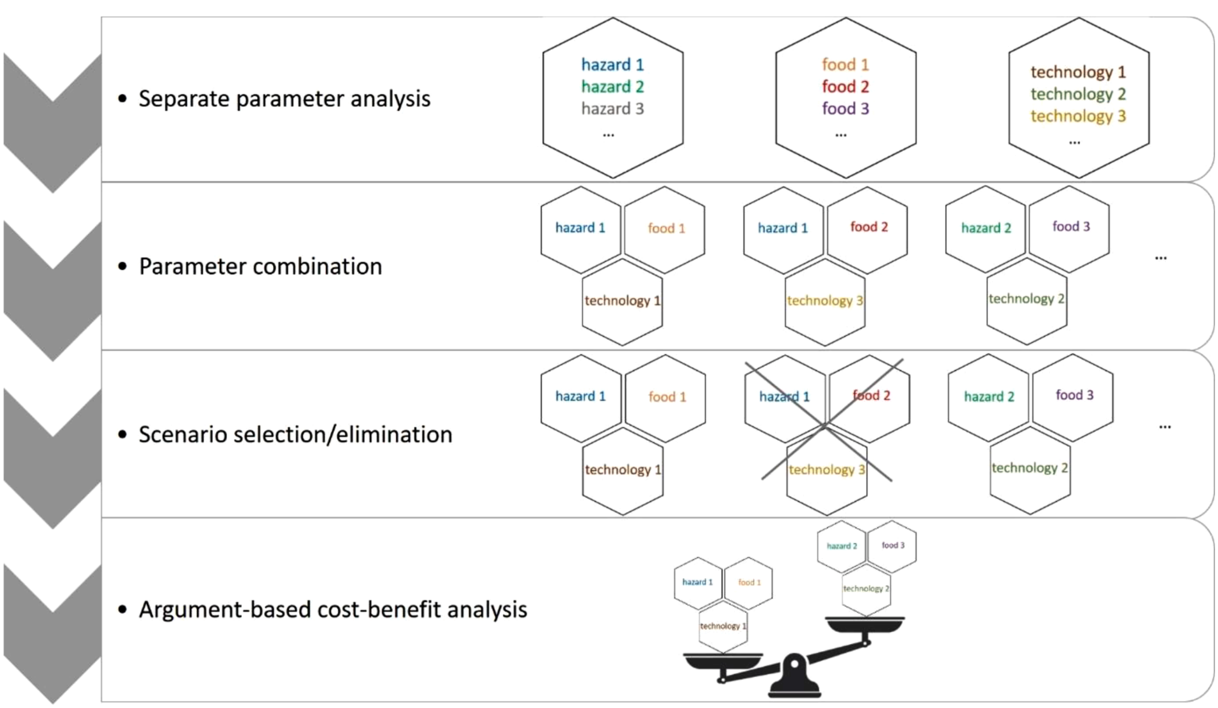

Fig. 2. Steps of cost-benefit analysis.

provision of shared folders where several users can share assets.

As part of Task 5.5 of the SAFFI project, the beta software will be presented and tested to end-users, who will then be tasked with using the software tool for a period of time. After this time, interviews and surveys will be conducted with the end-user sample who will provide feedback and recommendations from their variety of perspectives. This testing period allows for potential bugs in the software to be identified and for improvements to the software to be suggested. Following feedback from the end-users, an assessment will take place on the potential next steps to make the decision support system software more commercially ready and viable (Fig. 1).

2.2. Cost-benefit analysis

2.2.1. Objectives

Throughout the project and within its different research components, different parameters of food safety in the infant food chain are explored[15,16]). These include the hazards examined, the food products considered, the processing technologies used, the detection tools used for hazard control. Obviously, these parameters are not independent. A given hazard may be prevalent in some foods more than others, and hazards may be unequally impacted by a processing technology, to cite a few examples. Consequently, some scenarios,[10] that is to say, combinations defined by a given hazard in a given food product, undergoing a given processing technology and examined with a given detection tool, will be of salient interest for further discussion in the project. The first objective of cost-benefit analysis is to identify and select such scenarios. The second objective is to draw up a broad-view assessment involving various considerations[21] –from food safety, technical feasibility to nutritional interest, economic impacts, etc. – and various stakeholders[20] for these salient scenarios. The stakeholders involved in infant food safety control include consumers, who should benefit from health issue prevention; professionals of early childhood and healthcare, who play an advisory role with families; the food industry as well as public authorities, concerned about the prevention of health crises, the preservation of public confidence, the availability of efficient and affordable hazard detection tools, the creation of economic opportunities; researchers and industrials of food safety related technologies, etc. Stakeholders may bring different visions and different expectations from the research carried out.[3] The third and final objective is to conclude on the potential of the different scenarios, collectively and for each of the stakeholders and concerns examined.[4]

2.2.2. Methods

The steps of the methodology followed are depicted in Fig. 2.

In the first step, based on the knowledge available through a variety of sources for each parameter separately (hazard, food product, processing technology, detection tool), the pros and cons of focusing on each parameter value (a given contaminant, a given baby food, etc.) are determined and the information is structured and stored in the form of arguments. Argumentation, a reasoning model based on the construction and evaluation of interacting arguments, has been formalized in different disciplines including computer science and artificial intelligence[6,11] and adapted to various uses such as decision making.[2] Its interest in the food sector, together with other system-modeling approaches, has been underlined in several recent reviews.[12,19,1] Developments and argument structuration in the food sector can be consulted e.g. in Thomopoulos et al.[22, 23]

In the second step, all the possible combinations of parameter values are computed. The results are the “scenarios” considered.

Within the scenarios obtained, not all of them make sense. For instance, a processing technology may be irrelevant a given food product. In this case, all the scenarios where the incompatible processing technology and food product were combined can be removed. This is the object of the third step. If too many scenarios still remain, which may impair a thorough study of each of them, a careful selection of scenarios based on the project priorities may be relevant.

Finally, in the fourth step, the same approach as in the first step is applied at the scale of the scenario. To this end, the sets of arguments attached to the parameter values composing the scenario are merged and further completed by additional arguments proper to the combination defining the scenario. These arguments are elicited through multi-stakeholder discussions.

2.2.3. Outcomes

At the end of the fourth step, a so-called “collective attitude” measure can be computed for each of the scenarios considered, allowing the project consortium to compare them. Details on its exact computation can be found in Kurtz & Thomopoulos.[13] This measure can be computed in several modes: (i) either globally, for all stakeholders and concerns brought together, (ii) or modularly, for each stakeholder group or each concern separately. Mode (i) allows the project consortium to highlight the most consensual scenarios to implement and the underlying reasons, provided by the arguments associated with the scenarios. Mode (ii) allows the project consortium to highlight stakeholders or concerns that might most benefit from, or on the contrary be unsatisfied by some scenarios, and the underlying arguments. This is an essential point to anticipate the potential and possible risks of scenarios prior to their implementation.

3. Conclusion

Issues faced in the infant food chain in the EU and China are complex and require equally complex solutions to adequately enhance the reliability and transparency of food safety control. The SAFFI project proposes an integrative approach, essentially bridging the knowledge and data gaps that currently exist by integrating data and knowledge from a diverse range of sources, disciplines, stakeholders and actors. The decision support system, central to this integrative approach, will connect the expertise from disciplines including risk assessment, food technology, predictive toxicology, residue chemistry, predictive microbiology, pediatrics and knowledge engineering, the inputs of stakeholders and end-users and the wealth of food safety knowledge and data developed during the project to modern data science principles, and deliver an upgradeable and user-friendly tool. The proof of concept of this project will be exhibited in the application of the decision support system tool to four case studies in the infant food supply chain, and its success will be measured by the achievement of the scientific, technological, socio-economic, and regulatory objectives.

Declaration of Competing Interest

Authors declare no conflicts of interest.

Acknowledgments

The SAFFI project (Safe Food for Infant in the EU and China) has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement N°861917

References

1 Aceves Lara CA, Athès V, Buche P, Della Valle G, Farines V, Fonseca F, Guillard V, Kansou K, Kristiawan M, Monclus V, Mouret JR, Ndiaye A, Neveu P, Passot S, Pénicaud C, Sablayrolles JM, Salmon JM, Thomopoulos R, Trelea IC. The virtual food system: Innovative models and experiential feedback in technologies for winemaking, the cereals chain, food packaging and eco-designed starter production. *Innov Food Sci Emerg Technol.* 2018;46:54–64. https://doi.org/10.1016/j.ifset.2017.10.006.

2 Amgoud L, Prade H. Using arguments for making and explaining decisions. *Artif Intell.* 2009;173:413–436. https://doi.org/10.1016/j.artint.2008.11.006.

3 Bourguet JR, Thomopoulos R, Mugnier ML, Abécassis J. An artificial intelligence-based approach to deal with argumentation applied to food quality in a public health policy. *Expert Syst Appl.* 2013;40:4539–4546. https://doi.org/10.1016/j.eswa.2013.01.059.

4 Chaib RL, Macombe C, Thomopoulos R. Adaptation of a participatory system-modeling method to the constraints of remote working. In: *Proceedings of the Conference on Complex Systems - France (FRCCS 2021)*; 2021. May 2021. Easychair preprint n°6305, 20 pages https://easychair.org/publications/preprint/l8VF.
5 CremeGlobal, 2021. Data foundry [WWW Document]. URL https://www.cremeglobal.com/data-foundry/.
6 Dung PM. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artif Intell*. 1995; 77:321–357. https://doi.org/10.1016/0004-3702(94)00041-X.
7 European Commission, 2016. H2020 Programme. Guidelines on FAIR data management in horizon 2020 history of changes 12. Retrieved from: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf.
8 European Commission. 2015 Horizon 2020 - work programme 2014-2015. General Annexes: G. Technology readiness level (TRL https://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/annexes/h2020-wp1415-annex-g-trl_en.pdf.
9 European Commission, 2021. Monitoring EU agri-food trade: developments January-October 2021. https://ec.europa.eu/info/sites/default/files/food-farming-fisheries/trade/documents/monitoring-agri-food-trade_oct2021_en.pdf.
10 Godet, M., 2008. Strategic foresight: use and misuse of scenario building. LIPSOR Working Paper #10, 143 p. Retrieved from: https://prospectivayfuturo.files.wordpress.com/2017/09/godet-strategic-foresight-n10-2009.pdf.
11 Hunter, A., Besnard, P., 2008. Elements of argumentation. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 4724 LNAI, 4. 10.1007/978-3-540-75256-1_3.
12 Kansou K, Laurier W, Charalambides MN, Della-Valle G, Djekic I, Feyissa AH, Marra F, Thomopoulos R, Bredeweg B. Food modelling strategies and approaches for knowledge transfer. *Trends Food Sci Technol*. 2022;120:363–373. https://doi.org/10.1016/j.tifs.2022.01.021.
13 Kurtz A, Thomopoulos R. Safety vs. Sustainability concerns of infant food users: French results and European perspectives. *Sustainability*. 2021;13. https://doi.org/10.3390/su131810074.
14 Naumann F. Data profiling revisited. *SIGMOD Rec*. 2013;42:40–49. https://doi.org/10.1145/2590989.2590995.
15 Pettoello-Mantovani M, Mestrovic J, Namazova-Baranova MD, PhD L, Giardino I, Somekh E, Vural M. Ensuring safe food for infants: the importance of an integrated approach to monitor and reduce the risks of biological, chemical, and physical hazards. *J Pediatr*. 2021;229. https://doi.org/10.1016/j.jpeds.2020.10.074, 315-316.e2.
16 SAFFI Project (2020). Available online: https://www.saffi.eu.
17 Sprague RH. A framework for the development of decision support systems. *MIS Q Manag Inf Syst*. 1980;4:1–26.
18 Tao Q, Ding H, Wang H, Cui X. Application research: big data in food industry. *Foods*. 2021;10. https://doi.org/10.3390/foods10092203.
19 Thomopoulos R, Baudrit C, Boukhelifa N, Boutrou R, Buche P, Guichard E, Guillard V, Lutton E, Mirade PS, Ndiaye A, Perrot N, Taillandier F, Thomas-Danguin T, Tonda A. Multi-criteria reverse engineering for food: genesis and ongoing advances. *Food Eng Rev*. 2019;11:44–60. https://doi.org/10.1007/s12393-018-9186-x.
20 Thomopoulos R, Cufi J, Le Breton M. A generic software to support collective decision in food chains and in multi-stakeholder situations. In: *Proceedings of FoodSim 2020*; 2020:201–207, 6–10 September 2020 https://hal.archives-ouvertes.fr/hal-02484363v2/document.
21 Thomopoulos R. A practical application approach to argumentation for multicriteria analysis and decision support. *EURO J Decis Process*. 2018;6:237–255. https://doi.org/10.1007/s40070-018-0087-2.
22 Thomopoulos R, Moulin B, Bedoussac L. Supporting decision for environment-friendly practices in the agri-food sector: When argumentation and system dynamics simulation complete each other. *Int J Agric Environ Inf Syst*. 2018;9:1–21. https://doi.org/10.4018/IJAEIS.2018070101.
23 Thomopoulos R, Salliou N, Abreu C, Cohen V, Fouqueray T. Reduced meat consumption: from multicriteria argument modelling to agent-based social simulation. *Int J Food Stud*. 2021;10:133–149. https://doi.org/10.7455/IJFS/10.1.2021.A1.
24 VisualParadigm, 2022. Data Modeling: Conceptual vs Logical vs Physical Data Model [WWW Document]. Vis. Paradig. Online. URL https://online.visual-paradigm.com/knowledge/visual-modeling/conceptual-vs-logical-vs-physical-data-model#:~:text=The conceptual model is to,implementation of the data model. (accessed 3.1.22).