

## **Developing a framework for open and FAIR data management practices for next-generation risk- and benefit assessment of fish and seafood**

**Main author:** Javier Pineda-Pampliega (EFSA EU-FORA fellow)

**Co-authors:** Javier Pineda-Pampliega, Annetter Bernhard Annette, Rita Hannisdal, Ole Jakob Nøstbakken, Arne Duinker, Robin Ørmsrud, Froyland Livar, Gro-Ingunn Hemre, Gro Mathisen, Gisle Solstad, Josef Daniel Rasinger

### INTRODUCTION

The Norwegian Institute of Marine Research (IMR) collects samples of key marine species and receives samples from the Norwegian Food Safety Authority for national and international monitoring programmes. The IMR routinely reports data to the European Food Safety Authority (EFSA) and hosts an online tool – the Seafood Database (SF-DB) – through which a large curated in-house database can be accessed freely. With the OpenFoodTox 2.0 (OFT-DB) platform, EFSA publishes a comprehensive structured database in which hazard data collected by the agency in risk assessments is made available for download. The OFT-DB can be accessed from the EFSA Knowledge Junction hosted on the Zenodo platform, a project commissioned by the EC to support open data policy by providing an umbrella repository for EC-funded research. In the course of currently ongoing risk assessments performed at the Norwegian Scientific Committee for Food and Environment (VKM), a paucity of tools for easy integration and cross-referencing of information stored in the different databases was noted.

### METHODOLOGY

The objective of the present project, which is being conducted under the auspices of the EFSA EU-FORA programme (Series 5; 2021/2022), is to evaluate whether the SF-DB can be adapted for optimised data access and crosstalk with tools such as the OFT-DB. It will be assessed if, modelled on the EFSA OFT-DB, version-controlled downloadable copies of selected snapshots of aggregated data obtained from the SF-DB can be created and can be made available on the Zenodo platform. For a proof of concept, it will be tested if selected sub-sets of contaminants and nutrients of interest to ongoing risk-benefit assessments at the VKM can be extracted from the SF-DB. Using established and novel approaches to aggregate complete and left-censored datasets, it will be assessed if the extracted data can be statistically summarised to a degree of granularity suitable for further work and augmentation of SF-DB data with data from the EFSA OFT-DB and other databases relevant to the project.

## RESULTS

The results of this EU-FORA project are expected to provide version-controlled snapshots of combined occurrence and hazard data of selected contaminants relevant in relation to the consumption of fish and seafood. In addition, it will be evaluated whether the occurrence of specific nutrient data relevant for risk-benefit assessments of fish and seafood can be collated, aggregated and made available on Zenodo. Lastly, the work presented here will explore the suitability of publicly accessible repositories such as Zenodo for the dissemination of supplementary data accompanying risk assessment reports (such as outcomes of systematic literature searches, scripts used in data analyses and aggregated data extracts of different databases), which are commonly published in immutable formats that are not easy to re-use. At the conference, the experiences of the EU-FORA fellow with database setup and findable, accessible, interoperable, and reusable (FAIR) open data policies and their application in risk assessments and risk communication will be presented.

## DISCUSSION

In this project, it will be evaluated whether selected version-controlled snapshots of occurrence and hazard data of compounds relevant in relation to the consumption of fish and seafood can be created and made publicly available. It is envisioned that this will allow for easy access to data and contribute to the further development of fish- and seafood-focused data analyses, pave the way for computational access to seafood-focused data, and possibly allow for future application of machine-aided data mining and in silico analyses. The EU-FORA fellow's work will also lay the foundations for further considerations that evaluate whether a VKM-specific knowledge base in Zenodo, in which structured data accompanying risk assessments of VKM, can be made available in future. Due to the data-richness of such material, these became valuable resources in their own right; i.e. in terms of the research data lifecycle, it is now anticipated that any archived data should be stored in formats that allow for easy re-use and re-collection for future risk and risk-benefit assessments.