eData: A format and toolset for FAIR monitoring data

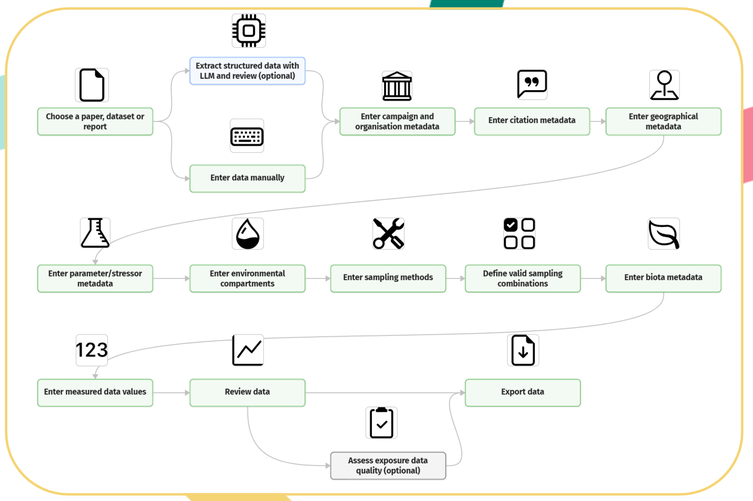
Sam A. Welch, Viviane Girardin & Knut Erik Tollefsen

Norwegian Institute for Water Research (NIVA), Økernveien 94, N-0579 OSLO, Norway.

Norsk institutt for vannforskning

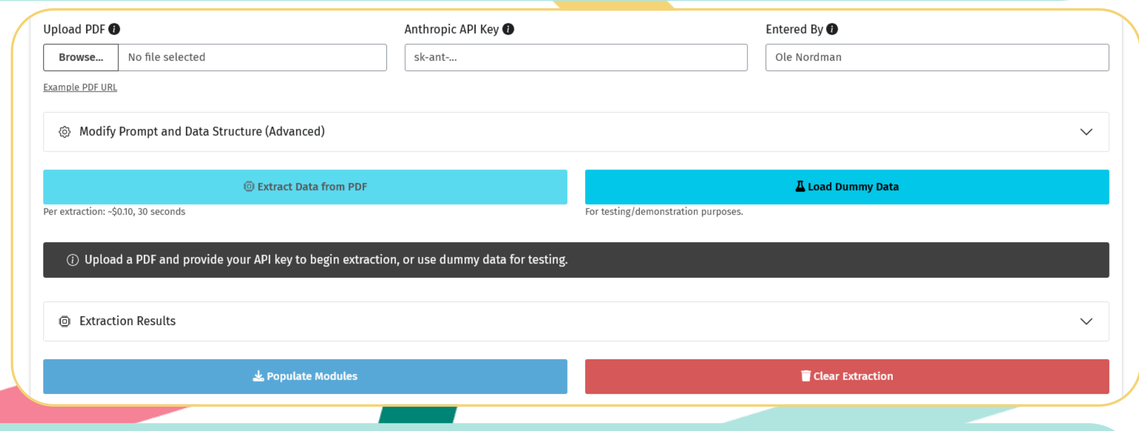
saw@niva.no

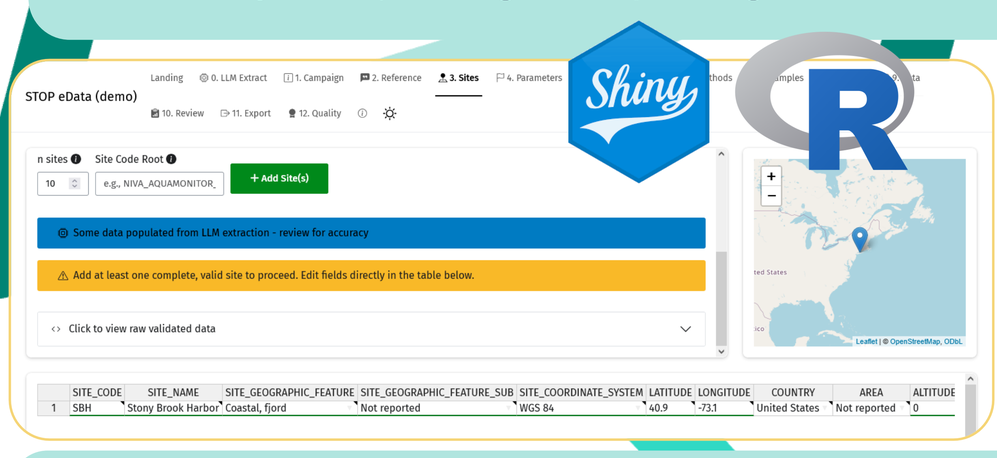
- **Good ecotoxicology requires good environmental monitoring data**, but we quickly forget the details of our data, making it harder to use well.
- The **FAIR principles** charge us with making our data **Findable, Accessible, Interoperable** and **Reusable** and so maximise value to data owners, the scientific community, and society — but making your data FAIR is yet another chore, and the benefits can be long-term and diffuse.
- We're developing **eData, a standardised format** and **interactive formatting tool** to help you **start the journey of FAIRifying your data** with the **minimum of upfront costs**.



- Differences in format, terminology, language, structure, and methodology make data significantly **less interoperable**.
- We're developing a **common, harmonised Data Reporting Format** (DRF), based on PARC P7.7.2, to capture biological, chemical and geographical exposure data.
- This makes data easier to **understand, compare, and synthesise, and use for risk assessment**.
- We're building tools to help us (and you) build evidence-based **Aggregate Exposure Pathways** for assessment

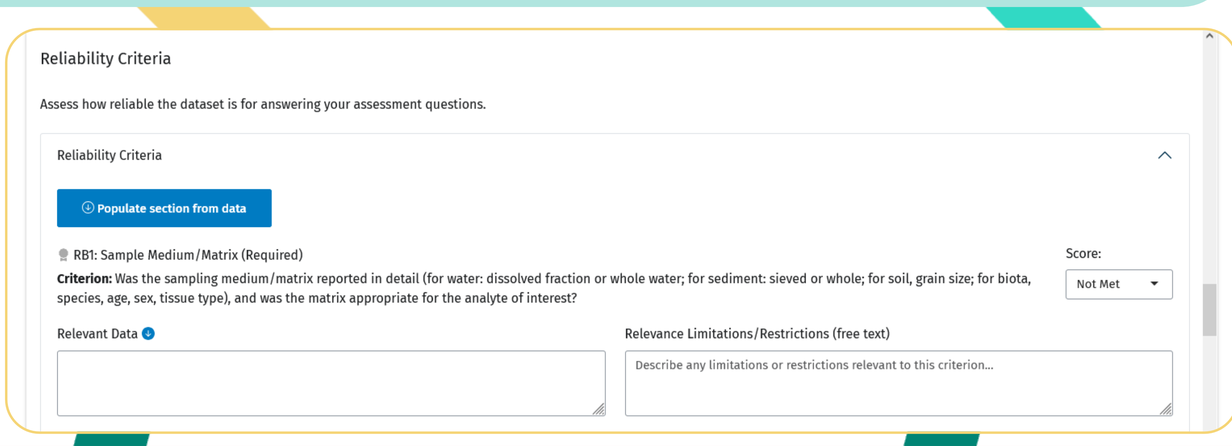
- Reformatting data by hand isn't **reproducible**, but code-based reformatting workflows can be **impractica**l.
- RShiny adds **interactive, modular app design** to the widely used programming language.
- This lets us optimise a data formatting workflow for **time efficiency** *and* **standardise use of language and format**.
- We need your help **testing and expanding** the toolset.



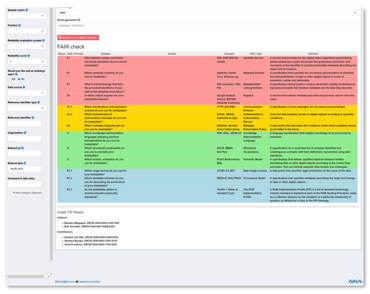


- **LLM-assisted data extraction** reduces the human workload, but can introduce **hallucinated data** (i.e. made up or inferred); human error also affects data quality
- A combination of **LLM-input, human sanity checking, and robust validation** and exploration tools **catch as many errors** as possible before they enter analysis and storage.

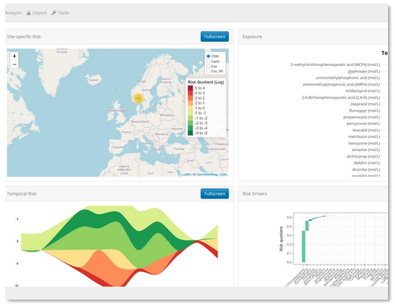
- We are also building an integration with the **Criteria for Reporting and Evaluating Exposure Datasets (CREED) guidelines** to allow quick assessment of a dataset's **relevance and reliability** in exposure or risk assessment.
- We aim to continue developing this module as part of an integrated network of **data-driven ecotoxicology tools** that will support rapid, efficient, reliable risk assessment and characterisation.



What's next after data formatting?



Archiving
FAIR can be intimidating, so we're making a module to help upload to **Zenodo**.



Analysis
eData can easily be quickly explored with standardised tools, or combined with **qData** for ERA

Reuse
You still own your data but reuse by you or others is faster and easier.

Is it ready to use?

- **Our prototypes work** and we use them internally, but development, testing, and improvement is an involved process.
- Building the kind of **high-quality, user-friendly** solutions that **ecotoxicology deserves** requires testers and contributors.
- We're always **looking for collaborations** and ways to help us help you — **get in touch at saw@niva.no**!

Find out more!

PARC Partnership for the Assessment of Risks from Chemicals

EXPECT

Co-funded by the European Union

This work was carried out in the framework of the European Partnership for the Assessment of Risks from Chemicals (PARC) and has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101057014, Research Council of Norway project EXPECT (RCN-315969) and NIVA's Computational Toxicology Program, NCTP (RCN-342628).

Gratefully Acknowledged
NIVA: Walter Zobl, Sophie Mentzel, Anna Hatland, Benno Dillinger, Li Xie, Jemmima Knight, Kim Leirvik, Andrea Merlina
PARC: T7.2.2 Team
MD: Dag Rosland
Icons: Angger Prasojo, Ida Desi Mariana, Freepik

Citations

Teeguarden, J.G., et al., 2016. **Completing the link between exposure science and toxicology for improved environmental health decision making: the aggregate exposure pathway framework**. Environ. Sci. Technol. https://doi.org/10.1021/acs.est.5b05311

Di Paolo, C., et al., 2024. **Implementation of the CREED approach for environmental assessments**. Integr. Environ. Assess. Manage. 20, 1019–1034. https://doi.org/10.1002/ieam.4909