

RDM @ Eawag

Scientific IT – Research Data Management

Structure



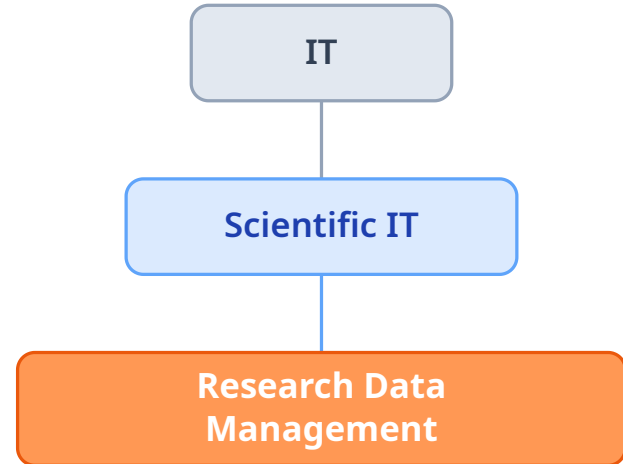
Stuart - Evolutionary Biologist turned HPC expert -
Team Lead



Sofya - Scientific Data Applications Specialist



Christian - Software Engineering and Operations with a
keen interest in Security



Support - Scientific IT

- **RDM**
- HPC support
- Infrastructure
- Consultation on best practices and other topics in:
 - Software Engineering (language agnostic)
 - Cybersecurity
 - ML/AI/Agents
- Regular open sessions

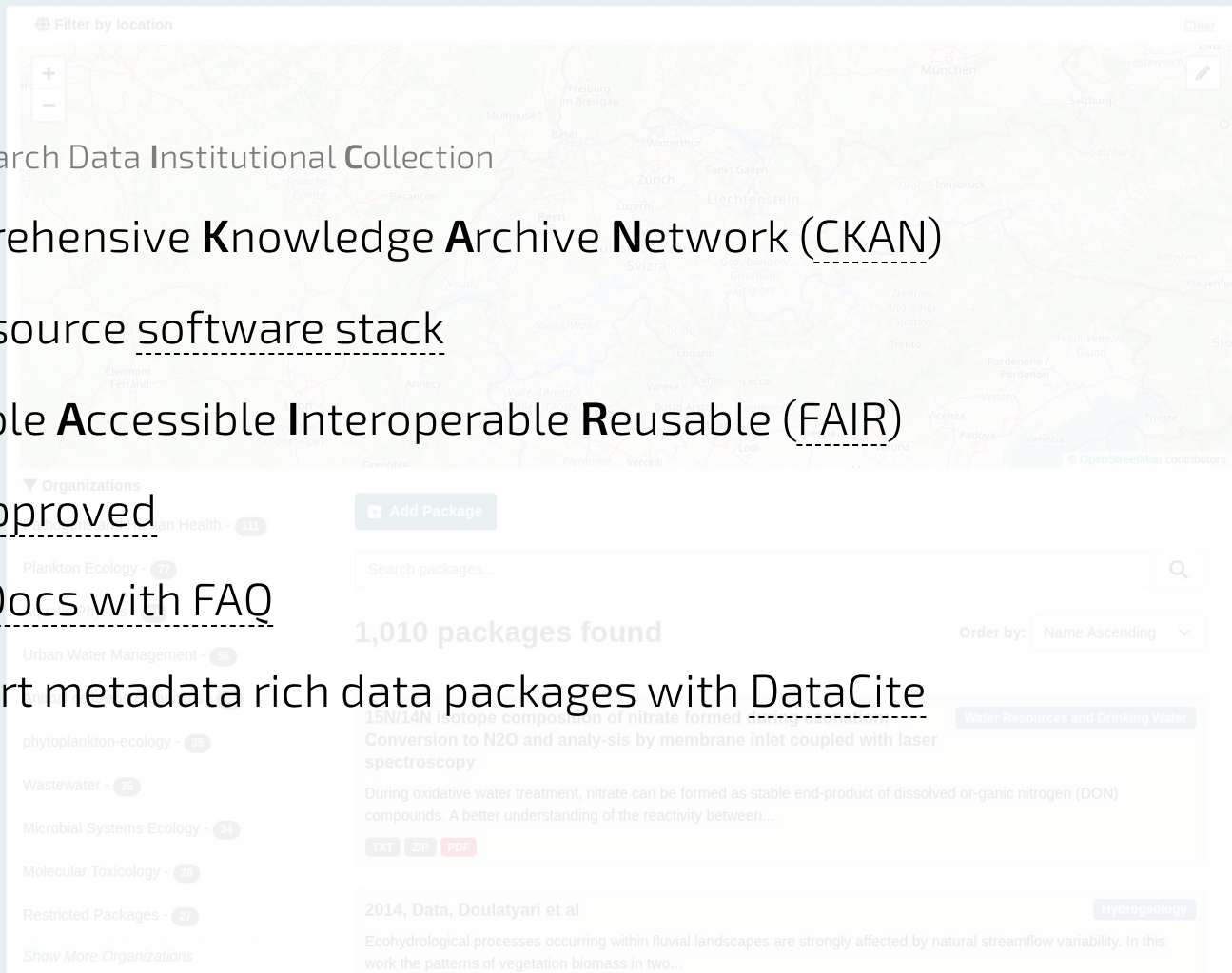
Support - RDM

- Planning (DMPs)
- Data transmission/transport
- Backup strategy, data safety
- Data security, encryption
- Legal issues, licensing
- Personal data, anonymization
- Dedicated storage (databases / network drive)
- Data Preservation
- Data Publication
- Data Quality checks
- DOI minting
- Metadata dissemination

ERIC

Eawag Research Data Institutional Collection

- **Comprehensive Knowledge Archive Network (CKAN)**
- **Open source software stack**
- **Findable Accessible Interoperable Reusable (FAIR)**
- **SNF approved**
- **RDM Docs with FAQ**
- **Support metadata rich data packages with DataCite**



ERIC/internal

The screenshot shows the ERIC data package interface. The main title is "Data for: Geringste Konzentrationen – Grösste Wirkung". Below the title, there is a description in German: "In sechs kleinen bis mittelgroßen Fließgewässern wurden die für aquatische Organismen extrem toxischen Pyrethroid- und Organophosphat-Insektizide mittels einer Spezialanalyse bis in den Picogramm pro Liter Bereich quantifiziert. An fünf der sechs untersuchten Standorte überschritten die gemessenen Insektizidkonzentrationen regelmäßig chronische und zum Teil akute Qualitätskriterien und die chronische Mischungsrisiko-bewertung zeigte während 43-100% des Untersuchungszeitraums hohe Risiken für die Invertebratengemeinschaft an. Werden Pyrethroid- und Organophosphat-Insektizide nicht in die Beurteilung der Gewässerqualität miteinbezogen, kann das Gesamtrisiko für aquatische Organismen erheblich unterschätzt werden." There is a "Dataset extent" map showing the study area in Switzerland. Below the map, there are "Data and Resources" listed: README.txt, Pyrethroids2018.xlsx, and Pyrethroids2017.xlsx. A "Metadata" section includes an "Author" list with seven names and their email addresses.

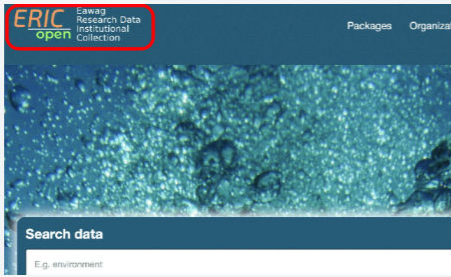
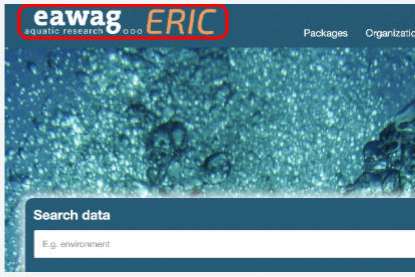
- Data preservation
- Support
 - Data organization, formats and annotation
 - Process automation
- No active review, quality control or curation
- Development & maintenance

ERIC/open

- **Data publication**
- DOI reservation & registration
- Support workflow in sync with article review & publishing
- (Meta)data dissemination, interlinking (ORCIDs, article-DOIs, ...)
- Review, quality control and curation
- Development & maintenance

The screenshot shows the ERIC/open interface for a data package. The header includes the ERIC/open logo and navigation links for Packages, Organizations, Projects, About, and Docs. The breadcrumb trail indicates the package is located under Organizations / Environmental Chemistry / Environmental Analytical Chemistry / Data for: Geringste... The main content area features a title 'Data for: Geringste Konzentrationen – Grösste Wirkung' with a DOI link (https://doi.org/10.2567/80001c7). A descriptive paragraph explains that six small to medium-sized fish species were exposed to pyrethroid and organophosphorus pesticides in a field experiment. A map titled 'Dataset extent' shows the study location in Switzerland. Below the map, there are links to 'Data and Resources' including README.txt, Pyrethroids2018.xlsx, and Pyrethroids2017.xlsx. A 'Citation' section provides the full citation for the data package. At the bottom, there are links for 'The associated article' and a list of associated articles.

ERIC/open vs ERIC/internal

	ERIC/open	ERIC/internal
Looks		
Link	opendata.eawag.ch	data.eawag.ch
Accessibility	World wide	⚠ Eawag intranet
Data Quality	Strictly checked	Unchecked
Interaction	Read-only portal	Read/write portal
Data	Immutable	Mutable
Use Cases	Access published data packages	Create/edit data packages

For the source of this comparison and other useful content on ERIC please look at our [docs](#).

Data Hygiene

Best practices throughout. Check out our minimal [guide](#) on publishing research data.

- Document everything (always include a README)
- Keep data well-organized from the start
- Use clear, consistent file naming (no spaces but "_" or "-", lowercase, standardized dates)
- Prefer open, simple, long-term file formats
- Respect legal and licensing constraints

Preservation Workflow

Getting started with ERIC:

- Log into ERIC/internal with your Eawag credentials
- Contact the ERIC data manager of your department for editor privileges

Upload to ERIC/internal:

- Create a package/group
- Upload your neat and organized data

Publication Workflow

Write to rdm@eawag.ch naming the package on ERIC you want us to publish.

RDM will get back to you with:

- A reserved **DOI** for your publication
- The data publication checklist for you to fill in and return
- Feedback on your data package in terms of data **quality, usability, protection**, a **basic code review** and more

Once all issues are resolved we can publish the data package.

Contents of your Data Publication

You should:

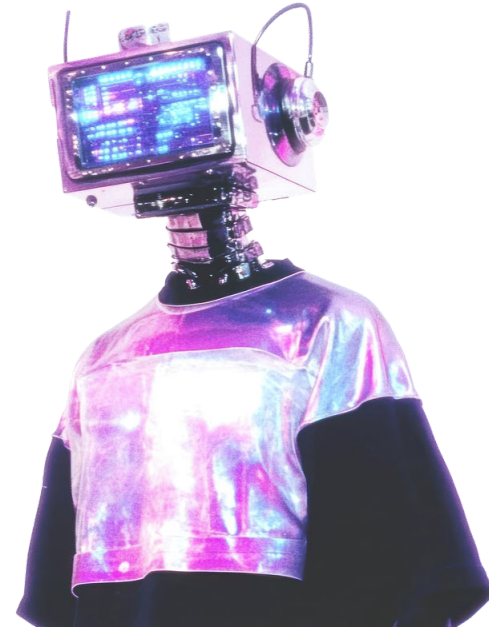
- Add your current code
- Add links to repositories, related packages
- Provide accurate and plentiful metadata

You should not:

- Add your paper to the data package
- Incorporate "Eawag specific information" or other sensitive data
- Violate data protection guidelines

Future

- Scientific-IT
 - New website coming in May
 - Flows for requesting services
- RDM
 - Streamlined support for DOI reservation and quality checks for scientists
 - Further data dissemination
 - Start planning a large upgrade for ERIC with the release of CKAN 3.0





Thank you! Questions?