Network of reference laboratories and related organisations for monitoring and bio-monitoring of emerging environmental pollutants

NORMAN digital storage

Archiving raw and environmental mass spectra for future analysis and assessment

Tobias Schulze, Werner Brack (UFZ)

Dübendorf – 17.09.2014

HELMHOLTZ CENTRE FOR ENVIRONMENTAL RESEARCH – UFZ

Retrospective analysis of water samples

Retrospective monitoring is the analysis of archived human or environmental samples regarding timely and spacial occurrence, trends and fate of priority and emerging compounds

How to keep water samples for retrospective analysis?

Environmental specimen banks

- Rothamsted soil bank since 1846 [1]
- Systematic institutional and governmental human and environmental ESBs since 1960s (e.g. Sweden, Japan, USA, Germany) [2]
- International Environmental Specimen Group since 2013
- Matrices (examples from German ESB)
  - Fish, muscles (riverine, marine)
  - Suspended particulate matter (Rhine, Saar, Danube, Elbe, Mulde)

But: No storage of water samples

[1] Jones et al. 1988, [2] http://www.inter-esb.org

Environmental specimen banks

- Long-term archiving of environmentally representative samples
- Standardised work flows for whole sampling and archiving procedure to yield homogenious and integrative samples
- QS/QC processes to ensure long-term stability and integrety of the samples
- Availabiltiy of samples to perform retrospective analysis depends on experience, standing and performance of the the laboratory

http://www.umweltprobenbank.de/en/

Long-term storage of raw water samples

Long-term storage of raw water samples is a challenge

- Homogeneity
- Contamination
- Stability (e.g. hydrolysis, biodegradation)
- Adsorption
- Filtration or not?
- Increasing volume of frozen samples
- …

Not recommended!

Archiving of water samples on solid phases

- Passive sampling using POCIS [1]
  - Good stability of solid-phases under cryo-conditions (-160 °C in gas phase of liquid N_2)
  - No information on stability of compounds

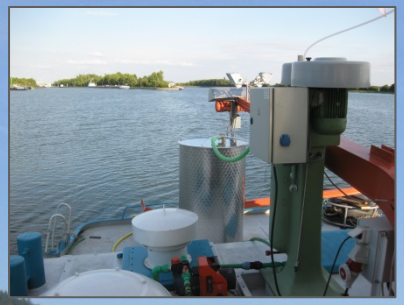
- Solid phase extraction (of large volumes) of water samples
  - Good long-term stability of compounds at low temperatures [2]

Open questions:

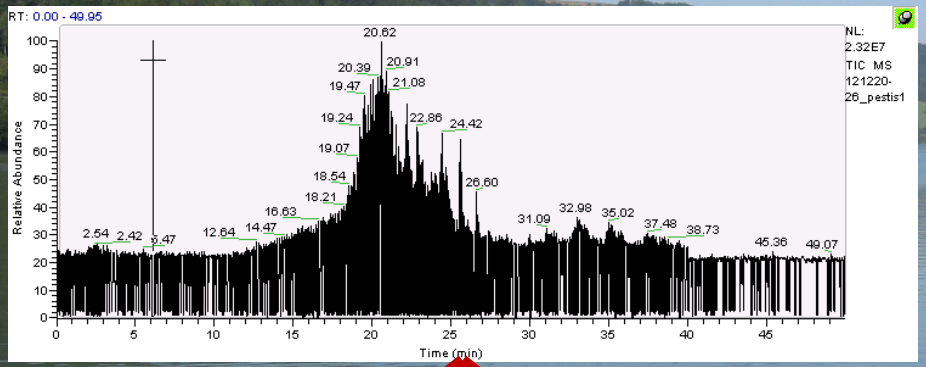
- Handling of SPE material (e.g. sticks to surfaces, homogeneity, freeze-drying)
- Representiveness (grab samples vs. integrative samples)
- Bias (selection of passive samplers and phases)

[1] Ricking et al. 2009; [2] Hillebrand et al. 2013

How to store water samples?

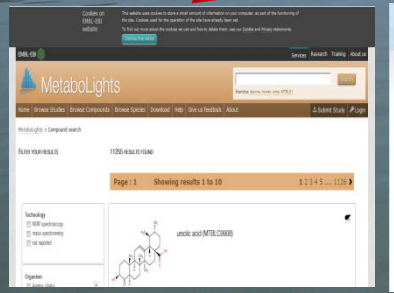


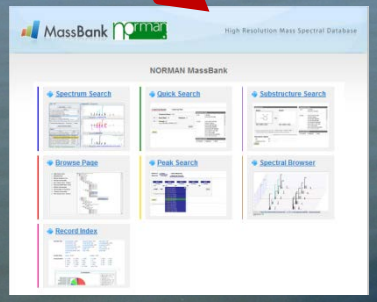
Storing of raw water samples is not recommended



HR-chromatograms of environmental samples could contain thousands of features

?





Usefullnes of storing all mass spectral information in repositories such as MassBank or „Envirolights“?

Sources: Andre Künzelmann (UFZ)
Tobias Schulze (UFZ)

The idea of the NORMAN digital storage is to archive whole chromatograms of water and other samples for retrospective analysis similar to ESBs in vendor‘s format

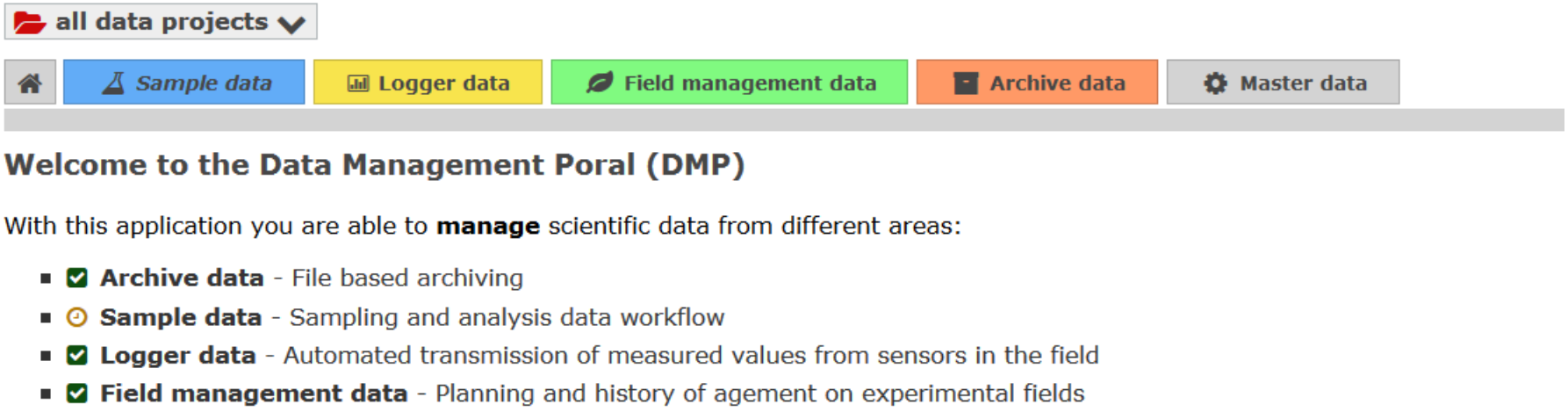
Why storing whole chromatograms?

- Software tools are in development (e.g. pick pickers)
- Important features might be not deconvolutet
- Exporting of chromatograms to open formats could yield in loss of information (wanted or unwanted)
- Users want to use raw data using sophisticated vendor‘s software for analysis

Requirements of a digital storage

- Archive storing files in long-term
- Support for large amounts of data (in the terabyte range)
- Redundancy to protect against data loss
- Verification of the stored data using checksums (integrity)
- Data description using a standardised metadata catalogue (e.g. Dublin Core, Inspire Core)
- Search and filter functions (e.g. location, matrix, river basin)
- Map view of dataset is an added value

Digital data storage at UFZ Leipzig



- Basic technology for storing large data files is available
- Long-term archive, redundant, standardised meta-data (Dublin core)
- Not implemented so far
  - Integration of sample data and archive
  - External user interface

Old external user interface

HELMHOLTZ CENTRE FOR ENVIRONMENTAL RESEARCH – UFZ

UFZ-DataLog
The data catalog of the UFZ

Contact | Imprint

German

Startpage My Data Search

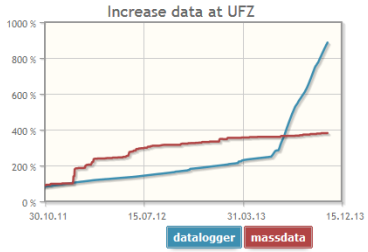
Login | Registration

Welcome to DataLog of the UFZ

Information about DataLog

In the course of time, the UFZ collects more and more data. These sometimes differ greatly from each other and are provided in different databases.

All these data have meta-information describing the single record. The aim of the DataLog is to help you search through this meta-information. By using this search, you can quickly search for specific data you need.



quick search

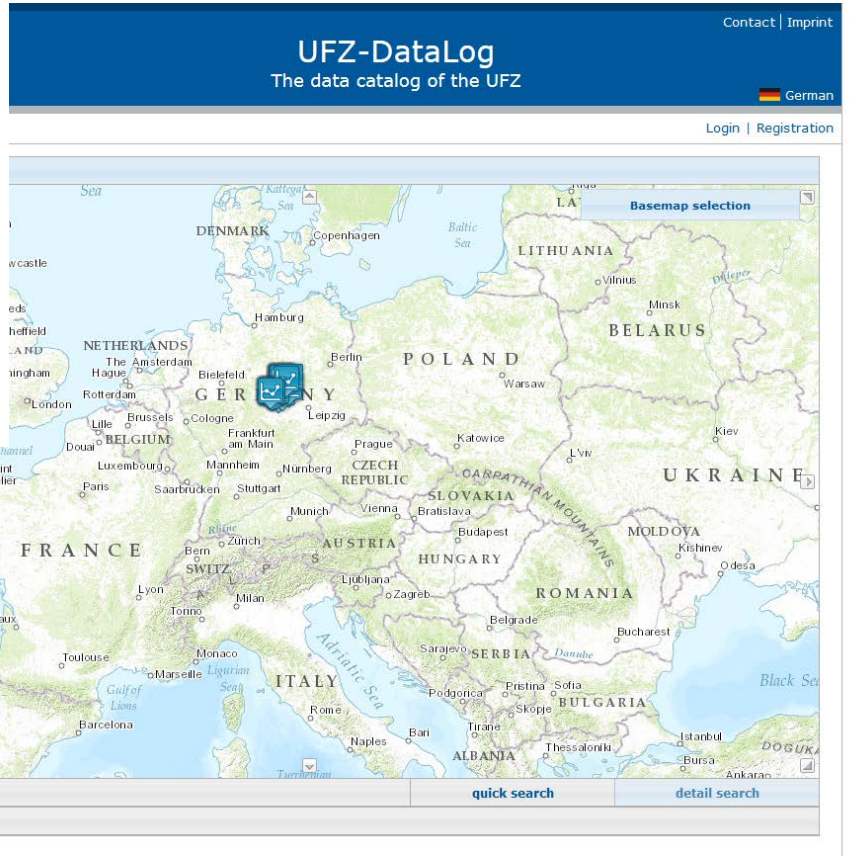
period of time

from

to

search

Atmosphäre **Bad Lauchstädt**
Bode-Selke CLE **Datenlogger**
Ermsleben **Halle** Hydrosphäre
Leipzig Magdeburg
Massendaten muldenberg
Pedosphäre smart TERENO
TO_Rappbode UFZ Wetterstation
Sauerbach YBW YTM



Requirements on the data submitted

- Chromatograms in raw format
- Additional vendors types files (e.g. reports)
- Additional information (e.g. spread sheets with evaluations in DCT format)
- Information on the methods used
- Data should be published under Creative Commons share alike license (CC BY SA)
- Metadata in standardised format (DCT, ISA-Tab, Inpire / Dublin core etc.)

Conclusions

- Long-term archiving is feasible
- Integrity of the data is warranted
- Coverage and quality of data is biased on methods / instrumentation
- Standardisation of data delivery
- Data accessible for the public without restrictions (CC BY SA)

"Opened up a Pandora's box" by F.S. Church (http://en.wikipedia.org)

Thank you for your attention!

Questions or comments?

Contact:
tobias.schulze@ufz.de