

NORMAN DATABASE WORKSHOP

Objective of the workshop

The main objective of the NORMAN Database Workshop is to discuss strategies for further development of the NORMAN Database System (NDS) and to provide a platform for technical discussions among the networks' IT experts. It is expected that the data science experts will discuss transition to next generation data science with the use of artificial intelligence, machine learning and advanced statistical and visualisation tools. An important issue will be a proposal of (i) a strategy to ensure long-term sustainability and resilience of the NDS and (ii) inclusion of the NDS as part of the European environmental data infrastructure.



HELLENIC REPUBLIC
National and Kapodistrian
University of Athens



INERIS

- When:** 26 Oct 2023 (09h00) – 27 Oct 2023 (14h30) - EET
- Where:** 'Ioannis Drakopoulos' conference room, Panepistimiou 30, Athens, Greece
- How:** Hybrid workshop - Registered participants will receive joining instructions by email
- Organizers:** Kapodostrian University of Athens, Environmental Institute and INERIS
- Contact person:** Dr. Jaroslav Slobodnik <slobodnik@ei.sk>
- Registration:** Open for NORMAN members

NORMAN DATABASE WORKSHOP

AGENDA

26 October 2023 – Day 1

08:30-09:00 Registration and coffee

NORMAN Database Workshop

09:00-09:15 Welcome by the host organisation, workshop organisers and outline of the workshop goals

Session 1 – Where are we? What is new?

09:15-09:45 NDS – a platform of interlinked database modules supported by tools for automated data processing (Jaroslav Slobodnik, EI)

09:45-10:30 What is new?

- NORMAN Digital Sample Freezing Platform and EMPODAT SUSPECT (Nikiforos Alygizakis, EI, 10 min)
- MassBank Europe and RMassBank (Tobias Schulze, UFZ/UBA, 5 min)
- Antibiotic Resistance Bacteria and Genes Database (Lian Lundy, LTU, 5 min)
- NORMAN BioActivity Database (Tessa Pronk, KWR, 5 min)
- Modes of biological action for monitored chemicals – where does it belong in the NDS? (Wibke Busch, UFZ, 5 min)
- **Proposals for better interlinking of the existing and new NDS modules – all participants** (Discussion)

10:30-11:00 Coffee break

Session 2 – Data quality and data traceability - novel strategies for data curation

11:00-12:30 Automated data curation tools

- The current workflow for removal of duplicities and sustainable update of the NDS; how machine learning tools could be of help to improve the quality of NDS data (Lubos Cirka, EI, 10 min)
- Application of machine learning approaches to the NORMAN Database System (Jana Schor, UFZ, 5 min)
- **Proposals of strategies for data curation and merging - open software/ packages/ approaches – all participants** (Discussion)

12:30-14:00 Lunch

Session 3 – From data to information – towards embedding artificial intelligence and machine learning tools in the NDS

14:00-15:30 Advanced tools in support of risk assessment of chemicals in the environment

- Automated tools for prioritisation of target and suspect substances in environmental matrices; integration of a Mixture Risk Indicator (Valeria Dulio, INERIS, 10 min)
- Strategy for sustainable update of KEMI Hazard and Exposure Scores based on REACH database (Stellan Fischer, KEMI, 5 min)

When: 26 Oct 2023 (09h00) – 27 Oct 2023 (14h30) - EET
Where: 'Ioannis Drakopoulos' conference room, Panepistimiou 30, Athens, Greece
How: Hybrid workshop - Registered participants will receive joining instructions by email
Organizers: Kapodostrian University of Athens, Environmental Institute and INERIS
Contact person: Dr. Jaroslav Slobodnik <slobodnik@ei.sk>

NORMAN DATABASE WORKSHOP

- NDS in support of Early Warning System for Europe (EWS PARC; Lutz Ahrens, 5 min)
- Overview of visualisation tools used in the NDS and inspired database systems; and
- Neural network-based statistical tools for source tracking of chemicals (Nikiforos Alygizakis, EI, 5 min)
- **Proposals of 'fit for purpose' machine learning and visualisation tools to be embedded in the NDS – all participants** (Discussion)

15:30-16:00 Coffee break

16:00-17:00 Breakout groups (parallel sessions)

- 1) What do we want to improve in the NDS?
 - Automated retrieval and update of data from external data sources
 - Upgrading the existing features of the NDS
 - Interlinking the modules, development of new modules
 - Visualisation features
 - Integration of data from models (e.g., the VEGA Hub)
- 2) What is our vision for implementation of artificial intelligence and machine learning tools in the NDS?
 - IT solutions for further upgrade of the NDS
 - Proposals for the systematic use of AI, machine learning, statistical and data visualisation tools

17:00-17:15 Coffee break

17:15-18:00 Breakout groups (cont.)

Group 1) and Group 2) exchange

20:00 Dinner (for those who have registered)

27 October 2023 – Day 2

09:00-09:20 Reporting from the breakout groups in Day 1

Session 4 – How to become fully FAIR?

- 09:20-10:00** **NDS as part of the European and global environmental data infrastructure; are the NDS data 'Findable, Accessible, Interoperable, and Reusable' by regulators and the scientific community?**
- FAIR Data Principles – an Overview (Christina Kratz, UBA, 10 min)
 - NDS inspired databases: Black Sea Environmental Data Platform (BS e-DataPlatform), LIFE APEX Database System incorporating OSPAR CONNECT and HELCOM PreEMPT data – existing and potential links to IPCHEM, ICPDR Water Quality Database, ICES, AMAP and EMODNet databases (Jaroslav Slobodnik, EI, 10 min)
 - Example of established API links between the NDS and national databases (UBA, 5 min)
 - **Proposals for interlinking the NDS (automated retrieval/exchange of data) with relevant external database systems – all participants**
 - **Proposals for making the NDS a part of the European and global environmental data infrastructure – all participants** (Discussion)

10:00-11:30 Coffee break

11:30-13:30 Plenary discussion

When: 26 Oct 2023 (09h00) – 27 Oct 2023 (14h30) - EET
Where: 'Ioannis Drakopoulos' conference room, Panepistimiou 30, Athens, Greece
How: Hybrid workshop - Registered participants will receive joining instructions by email
Organizers: Kapodostrian University of Athens, Environmental Institute and INERIS
Contact person: Dr. Jaroslav Slobodnik <slobodnik@ei.sk>

NORMAN DATABASE WORKSHOP

- Continuation of the discussion from Session 4
- How to ensure sustainability of the NDS and its support of the EU environmental policy and regulations?
- What are the partners we want to connect to?
- Where do we see the NDS in 5 years and what are the priorities?
- Final discussion and summary of take-home messages and further actions

13:30-14:30 Lunch

14:30 End of workshop

About the NORMAN Database System

The NDS is a joint activity of all NORMAN members and at the core of the NORMAN activities, providing data and tools to fulfil its goals and visions. The NDS consists nowadays of 13 integrated databases modules:

1. Suspect List Exchange - <https://www.norman-network.com/nds/SLE/>
2. Substance Database - <https://www.norman-network.com/nds/susdat/>
3. Chemical Occurrence Data (EMPODAT) - <https://www.norman-network.com/nds/empodat/>
4. Ecotoxicology Database - <https://www.norman-network.com/nds/ecotox/>
5. Digital Sample Freezing Platform (DSFP) - <https://norman-data.net/Verification/>
6. Substance Factsheets - <https://www.norman-network.com/nds/factsheets/>
7. NORMAN MassBank - <https://massbank.eu/MassBank/>
8. Passive Sampling - <https://www.norman-network.com/nds/passive/>
9. Antibiotic Resistance Bacteria/Genes - <https://www.norman-network.com/nds/bacteria/>
10. SARS-CoV-2 in sewage - https://www.norman-network.com/nds/sars_cov_2/
11. Bioassays Monitoring Data - <https://www.norman-network.com/nds/bioassays/>
12. Indoor Environment - <https://www.norman-network.com/nds/indoor/>
13. Prioritisation - <https://www.norman-network.com/nds/prioritisation/>

All NDS modules can be searched either individually or starting from the module 'Search All Databases' (<https://www.norman-network.com/nds/common/>), where any substance from SusDat can be searched and displayed with all available data for this substance in any of the database modules. Two new modules, EMPODAT-SUSPECT and BIOACTIVITY Database, are under development.

An automated prioritisation module in the NDS is available for the target substances archived in EMPODAT. A prototype of the prioritisation module for suspect substances archived in EMPODAT-SUSPECT is currently being developed.

All participants are strongly encouraged to test functionality of all NDS modules prior to the workshop!

When: 26 Oct 2023 (09h00) – 27 Oct 2023 (14h30) - EET
Where: 'Ioannis Drakopoulos' conference room, Panepistimiou 30, Athens, Greece
How: Hybrid workshop - Registered participants will receive joining instructions by email
Organizers: Kapodostrian University of Athens, Environmental Institute and INERIS
Contact person: Dr. Jaroslav Slobodnik <slobodnik@ei.sk>