

## Proposals for NORMAN Joint Programme of Activities 2022

<b>Title</b>	<b>NORMAN Database System (NDS)</b>
<b>Type of activity</b>	Database maintenance and continuous update
<b>Leader</b>	EI
<b>Topic / activities</b>	<p><b>Background / Justification for the proposed activity:</b></p> <p>The NORMAN Database System (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:</p> <ol style="list-style-type: none"> <li>1. Suspect List Exchange - <a href="https://www.norman-network.com/nds/SLE/">https://www.norman-network.com/nds/SLE/</a></li> <li>2. Substance Database - <a href="https://www.norman-network.com/nds/susdat/">https://www.norman-network.com/nds/susdat/</a></li> <li>3. Chemical Occurrence Data (EMPODAT) - <a href="https://www.norman-network.com/nds/empodat/">https://www.norman-network.com/nds/empodat/</a></li> <li>4. Ecotoxicology Database - <a href="https://www.norman-network.com/nds/ecotox/">https://www.norman-network.com/nds/ecotox/</a></li> <li>5. Digital Sample Freezing Platform (DSFP) - <a href="https://norman-data.net/Verification/">https://norman-data.net/Verification/</a></li> <li>6. Substance Factsheets - <a href="https://www.norman-network.com/nds/factsheets/">https://www.norman-network.com/nds/factsheets/</a></li> <li>7. NORMAN MassBank - <a href="https://massbank.eu/MassBank/">https://massbank.eu/MassBank/</a></li> <li>8. Passive Sampling - <a href="https://www.norman-network.com/nds/passive/">https://www.norman-network.com/nds/passive/</a></li> <li>9. Antibiotic Resistance Bacteria/Genes - <a href="https://www.norman-network.com/nds/bacteria/">https://www.norman-network.com/nds/bacteria/</a></li> <li>10. SARS-CoV-2 in sewage - <a href="https://www.norman-network.com/nds/sars_cov_2/">https://www.norman-network.com/nds/sars_cov_2/</a></li> <li>11. Bioassays Monitoring Data - <a href="https://www.norman-network.com/nds/bioassays/">https://www.norman-network.com/nds/bioassays/</a></li> <li>12. Indoor Environment - <a href="https://www.norman-network.com/nds/indoor/">https://www.norman-network.com/nds/indoor/</a></li> <li>13. Prioritisation - <a href="https://www.norman-network.com/nds/prioritisation/">https://www.norman-network.com/nds/prioritisation/</a></li> </ol> <p>A new module EMPODAT-SUSPECT (<a href="https://www.norman-network.com/nds/suspect/">https://www.norman-network.com/nds/suspect/</a>) has been developed in 2021 in support of the WG1 Prioritisation and already tested for functionality with a first dataset of 6.5 million of data entries.</p> <p>All NDS modules can be searched either individually or starting from the module 'Search All Databases' (<a href="https://www.norman-network.com/nds/common/">https://www.norman-network.com/nds/common/</a>), where a presence of any substance from SusDat in any of the database modules is shown with all existing data.</p> <p>Automated prioritisation modules in the NDS have been developed for the target substances and suspect substances, following the instructions and guidance from the WG1. The prioritisation modules for target substances have already been tested for several countries (e.g. France, Netherlands, Slovakia, Montenegro, Ukraine...), river basins (e.g. Danube, 14 countries, Dnieper, Dniester, Severskiy Donets...), sea regions (Black Sea, North-East Atlantic, Baltic Sea), polar regions (Antarctica). A database fully compatible with the NDS and a parallel automated prioritisation module for top predators and their prey has been developed within the EU LIFE APEX project (<a href="https://lifeapex.eu/">https://lifeapex.eu/</a>; ca. 800,000 data entries) specifically for biota samples and will be integrated with the NDS in 2022.</p> <p>Each of the NDS modules is being continuously updated, either in terms of the required structure of the data (Data Collection Templates – DCTs downloadable at each of the modules) or data. The EMPODAT database grew from ca. 11 million to ca. 20 million data entries in 2021. The Netherlands and the UK Environment Agency were providing data to support their national prioritisation schemes required under the Water Framework Directive. A capacity of SusDat to support suspect screening has been increased from ca. 65,000 substances to ca. 90.000 substances in each sample stored in DSFP (&gt;2,500). A model-derived PBT properties were added to ca. 65,000 SusDat compounds in the Substance Factsheets module. An exposure score of compounds from the REACH database, converting confidential information on their production volumes and widespread of use into a score (0-1) is now available for ca. 98,000 substances. Predicted ecotoxicological threshold values in freshwater (fw), marine water (mw), biota (fw/mw) and sediments (fw/mw) are now available for ca. 65,000 compounds. The Substance Factsheet database is updated each three months to reflect latest changes in the US EPA CompTox Chemicals Dashboard database. A process to visualise the data in individual NDS modules is on-going.</p> <p><b>Description of the proposed activity and expected outcomes for 2022 (and beyond):</b></p> <p>The maintenance of the NDS and its continuous upgrade for new data takes a significant effort. The tasks for 2022 include:</p> <ul style="list-style-type: none"> <li>• Interlinking all NDS modules and quality check of all input data;</li> <li>• NDS Chemical Occurrence Data (EMPODAT): maintenance, upgrading and feeding of new data into the database; sharing the data with IPCHEM;</li> <li>• Continuous upgrade of all DCTs for an extended list of NORMAN substances (SusDat), drop-down lists and definitions of obligatory parameters;</li> <li>• Further upgrade of automated quality control tools for identification/ removal/ flagging of outliers in the collected datasets in EMPODAT;</li> <li>• Further development of data mining tools to extract raw data from IPCHEM and other database systems and establishment of a workflow for their processing into the 'NORMAN format';</li> <li>• Enhancement of visualization and data analysis capabilities of NDS;</li> <li>• Continuous upgrade and maintenance of SusDat;</li> </ul>

	<ul style="list-style-type: none"> <li>• Update of Passive Sampling module with new datasets; design of Passive Sampling – SUSPECT module;</li> <li>• Upload of new data into the ARBs/ARGs module;</li> <li>• Upload of new data into the NORMAN SCORE Database - SARS-CoV-2 in sewage (SC2S);</li> <li>• Upgrade of Substance Factsheets module – systematic collection of all data needed for prioritisation and data download functions;</li> <li>• WG2 – Upload of results of bioassays analyses from various projects;</li> <li>• WG3 and WG5 – Upload of results from various monitoring projects;</li> <li>• WG6 – Upload of data from the ILS on Indoor Environment 2020 – 2021;</li> <li>• CWGA PS – Upload of data from ILS 2021;</li> <li>• Collaborative trial NTS in biota - Upload of data from the ILS 2020 – 2021;</li> <li>• Upload of data from wide-scope target screening and suspect screening of chemicals in top predators and their prey carried out in the LIFE APEX project;</li> <li>• Upload of data from OSPAR CONNECT and UBA-HELCOM projects.</li> </ul> <p><b>Added value / Link with other NORMAN activities and / or other projects</b></p> <p>The proposed tasks will benefit all WGs and CWGAs in the NORMAN network.</p>
<b>Participants</b>	EI, all interested members
<b>Proposed in-kind contribution</b>	All – contribution of existing data EI – overall coordination
<b>Contribution needed from NORMAN Association<sup>1</sup></b>	Maintenance and continuous update of the NDS: - EI: 27,000 € Rental of the server hosting the NDS and backup system: - EI: 6,600 €

<sup>1</sup> Please, provide here a transparent justification of the requested resources and of the in-kind contribution, thereby distinguishing between the costs associated with “person-months” for the organisation, the “travelling costs” for invited speakers and the costs for the logistics (e.g. meals, room rental etc.)