

Proposals for NORMAN Joint Programme of Activities 2025

Title	NORMAN WG-1 Prioritisation
Type of activity	Working Group activities
Leader	Valeria Dulio (INERIS), Peter von der Ohe (UBA), Nikiforos Alygizakis (EI)
Topic / activities	<p>Background / Justification for the proposed activity:</p> <p>Better chemical management in line with the goals of the EU Chemicals Strategy and Zero Pollution Action Plan strongly relies on early identification and control of hazards and pressures through efficient monitoring and knowledge of the properties and use patterns of chemicals. Given the high number of chemicals present in our environment, prioritisation of chemical contaminants is highly demanded by regulators and decision-makers in order to identify and justify the most urgent actions necessary to achieve these goals.</p> <p>The tasks of WG-1 aim to:</p> <ul style="list-style-type: none"> ▪ Ensure that all relevant information for assessment of chemical risks is maintained and regularly updated in a timely manner; ▪ Integrate and exploit this information to identify lists of chemicals in need of priority actions (each priority list corresponds to an action); ▪ Continuously upgrade the prioritisation scheme (integration of innovative techniques) to enhance the value of the prioritization work and ensure a more reliable and meaningful interpretation of the results, particularly for the early identification of warning signals. <p>Description of the proposed activities and expected outcomes for 2025 (and beyond):</p> <p>Task 1: Support the prioritisation work of the Commission services at European level and provide comment on relevant consultation documents (PARC, EWS, where relevant) (ALL)</p> <p>NORMAN participates as a stakeholder in the WG Chemicals of DG ENV and intends to further contribute to the activities related with the review of the list of WFD Priority Substances and the Watch List.</p> <p>NORMAN WG-1 is committed to collaborate actively with PARC partnership on the following prioritisation-related topics:</p> <ul style="list-style-type: none"> ▪ Collaboration in the development and implementation of the mechanism for priority setting in environmental and multi-source monitoring – <i>MonitoringFrame</i> project (PARC T4.2.1). The NORMAN scheme was recognised as an efficient and pragmatic way to address regulatory questions while addressing existing data gaps. The concept was further developed in the Prioritisation tool designed within the <i>MonitoringFrame</i> project. Further collaborative actions remain open for discussion. ▪ Application of the prioritisation workflow to the list of candidate EDCs provided by PARC (collaboration WG-1 and PARC). A comprehensive compilation of information on regulated and non-regulated potential EDCs has been performed in PARC T4.2 in 2023. The list has been submitted for curation and inclusion in the SLE. Further to screening of the data collections available in the NDS, it will be possible to prioritise the compounds using the NORMAN scheme. ▪ <i>Early Warning System for Europe</i> (EWS): NORMAN already contributes by providing signals obtained from retrospective suspect screening of HRMS data (DSFP) and target monitoring data (EMPODAT), using indicators customised for the identification of early warning signals. NORMAN can also contribute with experimental and predicted data of hazardous properties as well as identification potential risks and observed impacts from research papers and other public sources (see Task 10). <p>Task 2: Collection and prediction of compound-specific information in support of prioritisation (UBA / EI / NKUA / DERAC / INERIS) (ongoing activity to be pursued in 2025)</p> <p>After finalisation of the harmonised ecotox DCTs, the still pending extraction and compilation of additional experimental ecotoxicity data from model training sets, will be started in 2025 (see also task 3). Regulatory quality targets for various matrices, including re-use, soil and marine waters, will also be collected and compiled (see also task 3).</p> <p>Moreover, data on physico-chemical and fate properties (e.g. Kow, Koc, BCF), as well as hazardous classifications (i.e. related to ED, CMR, PBT, PMT) will be retrieved e.g. from REACH registration dossiers and future CLP dossiers and integrated in the Substance Factsheets to support the hazard score in the prioritisation module (see task 4). For this purpose, a specific DCT has been developed in 2024. In case of lack of experimental data on these hazards, prediction models using artificial intelligence will be developed and used to fill the gaps.</p> <p>The P, B, T, CMR and ED scores were assigned to 76,269 compounds in the SusDat list by UBA and EI using the JANUS model. However, difficulties to run the model for the remaining ca. 30,500 chemicals persist. Hence, it is proposed to develop a quantitative model for persistence, based on a set of 8000 reported degradation half-lives for the OECD 302 tests. The models will allow to predict the DT50 in water for all compounds in SUSDAT that are covered by the model domain.</p>

In order to ensure that only reliable predictions are reported in the Substance Factsheets and used for the prioritisation of chemicals which lack experimental data, more general actions are planned as part of a new CWG focused on Data Science for Monitoring of Emerging Environmental Substances. Here **experts in modelling and machine learning should work to reach consensus on the use of the predictive models to be used and rules to define their application domain.**

Task 3: Collection of existing PNECs and deep learning-based toxicity predictions for use in NORMAN PNECs (UBA / EcotoxCentre / DERAC / EI) (ongoing activity to be pursued in 2025)

After the successful extension of prediction models for 3 additional fish species, 2 crustaceans and one insect species for **acute effects**, we propose to **extend the models also to chronic effects in fish, daphnia and algae for aquatic species with sufficiently large experimental datasets**. The final aim is to enable a more robust statistical approach for the derivation of freshwater P-PNECs and their subsequent conversion for application to other matrices, like sediments and biota, e.g. by using Koc and BCF (see task 2). Thereby, the use of chronic-based endpoints will allow to lower the AF for deriving the P-PNEC, which will be appreciated by many regulatory bodies that are using the prioritisation results of NORMAN.

For 2025, it is therefore planned to **continue with the predictions of toxicity values for both existing compounds and those that have recently been added to SusDat**, in order to be up to date with the list of Lowest PNECs. The upload of existing PNEC and the derivation of new PNECs will also include **MAC-EQS values**. This will allow the evaluation of peak exposure concentrations, such as those resulting from the German small water body monitoring project.

It is also planned to derive more robust **PNEC marine values** to support prioritisation of compounds in the **marine environment** (delayed from JPA 2024).

WG-1 will continue to **promote and coordinate the participation of Ecotoxicity Experts to derive and approve (i.e. vote for) new or revised Lowest PNEC values** for substances of the SusDat list, with a specific focus on substances that were highly prioritised in EU projects using the NORMAN Prioritisation Framework. The aim is to **progressively replace predicted PNEC values for substances prioritised in Cat 3 and 5, by experimentally-based PNEC values**.

Task 4: Consolidation of the Phys-Chem and Hazard properties module and integration in the NORMAN Database System and Prioritisation Tool (UBA, INERIS, EI in collaboration with WG-1 partners)

Physico-chemical and hazard properties are essential for deriving the Hazard scores (PBT/vPvB, PMT/vPvM, CMR, ED) used in the NORMAN prioritisation framework. Data are already collected in the NDS Factsheets for various relevant parameters (Kow, Koc, BCF, DT50, etc.), mainly from US EPA CompTox DashBoard.

An increasing number of models (e.g. Janus), including machine learning (ML)-based models (e.g. ToxAI), and tools are being developed to assess the Persistence (P), Bioaccumulation (B), Mobility (M), and Toxicity (T) criteria and classify the substances in line with the REACH and CLP requirements.

To ensure a comprehensive overview of available resources and enhance the assessment of data from various sources, it was decided to create a module with different sub-modules. These sub-modules will store both raw data (along with associated metadata in a harmonized format) and existing classifications. This module will also serve as a platform for the transparent assessment of data sources, facilitating consensus on robust classifications for prioritization purposes. Additionally, the module should allow for the selection of different options or scenarios (e.g., conservative, robust, average).

The design of the prototype started in 2024 (INERIS and UBA) and was discussed with the WG-1 participants. The work will be pursued in 2025.

The improvement of the features of this module will be coordinated by UBA and INERIS in consultation with the WG-1 partners and the programming will be performed by EI.

Task 5: Compilation of data / information regarding 'Use categories' and 'Chemical Functional Use' for all SusDat compounds (LCSB / INERIS / UBA / EI) (activity postponed and revised from 2024)

We propose to integrate a new "Use" module into the NDS to ensure that substances in SusDat will have suitable information about uses and functional categories. This information is crucial for the effective exploitation of priority substance lists and should be directly linked and searchable from the Substance Factsheets, SusDat and the Prioritisation tool. This new functionality will enable searching in the NDS by e.g., all pesticides, all herbicides; all pharmaceuticals, all antibiotics; or by, e.g. all PFAS, pyrethroids, PAHs, ionic surfactants.

The task proposed in the previous JPAs to address this objective remained on stand-by due to the complexity of reaching a consensus on creating a controlled vocabulary for use categories and functional classes. It is now proposed to adopt a data-driven approach, choosing a bottom-up rather than a top-down methodology, which will ensure that these tasks are suitable for a student but also deliver preliminary data upon which to develop appropriate controlled vocabularies.

The task will therefore involve:

- Pilot efforts to retrieve “Use” information from PubChem and other sources (e.g. Wikipedia, Wikidata, CompTox) and integrate it into a temporary “Use” module (working space).
- Organize the collected information to create a list of use categories and functional classes, along with their associated synonyms (ontology).
- Programming of a workflow resulting from the pilot efforts for the transfer of the information retrieved from PubChem and other sources into the actual prototype of the future “Use” module
- Integration and operational implementation of the prototype in the NDS (EI).

The task will involve:

- at least two on-site training sessions at LCSB to familiarise the team (LCSB, INERIS....) with the functionalities necessary for effectively exploiting the information available in PubChem.
- meetings with WG-1 participants to discuss the progress of the work and actions needed.

Task 6: Prioritisation framework: follow-up activities to integrate and test the new workflow as part of the Prioritisation Tool in the NDS (EI, INERIS, UBA, ALL) (ongoing activity to be pursued in 2025)

In 2024, critical steps have been taken for the development of the necessary infrastructure (back-end) to implement the online prioritisation tool. The chemical occurrence database now has a new EMPODAT API (https://www.norman-network.com/nds/empodat/api_matrix_index.php) that allows fast retrieval of target screening data. The DSFP has made publicly available the data schemas used in the platform (<https://dsfp.norman-data.eu/data-schema>). Among the schemas, the DSFP Detection schema allows for hosting detected suspected compounds, their IP scores and sub-scores, and the semi-quantified concentration levels in a flexible JSON file for a given compound and collection. As of November 2024, four collections have been screened for all SusDat compounds, formatted using the DSFP Detection schema, and are ready for data retrieval through the DSFP API (<https://server.norman-data.eu/search/detections/all>). Therefore, the infrastructure is ready, and it is possible to operationalize the link between target and wide-scope suspect screening investigations.

The focus of 2025 will be on:

- **Applying the Workflow:** Implementing the workflow published as a scientific article in 2024.
- **API adaptation (if necessary):** Making necessary modifications to the APIs to ensure they align with the workflow's requirements and purpose.
- **Enhancing Visualization (see Task 8):** Applying the workflow within the Prioritization module will produce more complex outputs, necessitating effective visualization. The visualization dashboard developed in 2024 will be further refined and fully integrated into the Prioritization module.
- **Quality Assurance:** Testing the workflow with the wastewater case study to validate the quality and reliability of the outputs.
- **Expanding Applications:** Extending the workflow's application to a different case study to demonstrate its versatility and robustness.

Task 7: Improvement of the features of the prioritisation tool (expert consultation and programming activities) (ALL) (EI, INERIS, UBA, UFZ, NKUA, OVAM, DERAC, LCSB, KWR)

The prioritization module provides as an output a table file. The integrated NTS prioritization scheme yields rich outputs and a table is not enough for comprehensive explanation of the output. More advanced data visualization and exploration tools are required for an in-depth analysis. In 2023 at the Database workshop in Athens, some visualisation options have been discussed and proposed for testing. In 2024, a visualization dashboard was developed (available at <https://norman-data.eu/Priportal>) and presented during the WP1 meeting (November 2024).

In 2025, the visualization dashboard will be further enhanced with additional features and tools, enabling users to perform a wide range of statistical operations more effectively. Enhancements in related tasks, such as the inclusion of use categories and information on tonnage bands, will create additional opportunities for further refinement. The dashboard will be finalized and integrated to the Prioritization tool. It will allow the users to draw more conclusions about the prioritisation results (within the action categories and beyond). The dashboard will be further leveraged to improve the visualization functionalities of other modules of the NDS.

Task 8: Prioritisation of data from 5th Joint Danube Survey 5 (collaboration with ICPDR) (INERIS, UBA, EI in collaboration with WG-1 partners) (end of 2025 – 2026):

Prioritisation of contaminants in the Danube River Basin with the following specific objectives:

- Identification of River Basin Specific Pollutants (update)
- Proposal for a candidate Watch List
- Changes / Progress observed from the previous Surveys

The work will be based on the extended NORMAN Prioritisation Framework using target and suspect screening data. Moreover, the improved PNECs (additional endpoints and species from ToxAI tool) will be applied.

Task 9: Testing state of the environment indicators in the Danube River Basin (NORMAN collaboration with 5th Joint Danube Survey) (EI, UBA, INERIS and ALL) (2025-2026)

The aim is to integrate the risks of individual substances at specific sites into an overall assessment (i.e. mixture toxicity) by mapping the concurrent presence of chemical contaminants across various

	<p>compartments in Europe. For example, determining "how many compounds exceed the lowest PNEC at each site" or assessing "the total mixture risk at each site." These approaches could serve as new chemical indicators to measure progress in environmental quality, e.g. as a result of remediation measures and inform policy decisions (cf. pesticides indicator of the EEA) about the trends. This would allow NORMAN to, e.g., support the Zero pollution monitoring report of the EEA. Moreover, analysis of datasets regarding the minimum (most toxic compounds) compared to the total effect (mixture risk) might allow to derive an indication of a retrospective mixture allocation factor (MAF) in support of the COM EGD initiative.</p> <p>Task 10: NORMAN contribution to Early Warning System (UBA, EI, INERIS, KWR, OVAM, EAWAG, SLU, NKUA)</p> <p>NORMAN has been working actively on EWS concept development for so many year on these topics, with NormaNEWS, DSFP, prioritisation scheme by action category, etc. and we believe that NORMAN can significantly contribute to the on-going collective efforts together with PARC in building an EWS for EU. NORMAN already contributes by providing signals obtained from retrospective suspect screening of HRMS data (DSFP) and target monitoring data (EMPODAT), using indicators customised for the identification of early warning signals.</p> <p>With this proposal WG-1 intends to develop a concept to support a European EWS by creating automated signals based on non-target and target screening data for compounds in the NDS, as well as through the development of a DCT to compile additional evidence from research papers and other public sources regarding emerging hazards (e.g. new effects), sources (e.g. new products), increasing exposure (e.g. first or increasing detection in DSFP) or observed impacts.</p> <p>These signals should be compiled and stored in a new EWS module. For the first time, an online form will allow experts to register their findings in a harmonized way. These "signals" can then be analysed by experts (for specific topics), or by the next generation of large language models (LLM), also known as General Artificial Intelligence (GAI), which are expected to emerge in 2025. Even if the next generation models are not yet able to link the multiple lines of evidence from the various signals submitted, NORMAN should start to compile the information needed to be ready once the GAI is.</p> <p>Task 11: Collaboration with other WGs to support prioritisation of contaminants in new compartments</p> <ul style="list-style-type: none"> ▪ Prioritisation of contaminants in soil (WG-1 and WG-7): (see JPA proposal by WG-7). ▪ Prioritisation of contaminants in the marine environment (WG-1 and WG-8): (see JPA proposal by WG-8) ▪ Prioritisation of contaminants in reused matrices (WG-1 and WG-5) (ongoing task to be pursued in 2025): In 2023, WG-5 revised the design of six EMPODAT DCTs to allow their application to reused matrices. New occurrence data based on literature review were submitted to EMPODAT for reused practices, i.e. stormwater, reclaimed water and sewage sludge in 2024. Moreover, new quality targets for reused matrices in agricultural practices were collected and will be uploaded in the ECOTOX database, and further quality targets in sewage sludge for agricultural fertilisation will be derived (Task 3 in WG-5). Thanks to this new input in 2024, it will be possible to start testing the NORMAN prioritisation workflow on the specificities of the reused matrices. In this case, the user interface will allow the user to select matrix and the reuse route, and the system will automatically select the proper PNECs, depending on the matrix and the prioritisation objective. For example, for prioritisation of contaminants of concern in reclaimed water for reuse in irrigation, it will be possible to select the wastewater effluent datasets labelled as "reclaimed water / reuse irrigation" and match the concentration data with the corresponding PNEC_crop, specific for protection of human health. <p>Added value / Link with other NORMAN activities and / or other projects</p> <ul style="list-style-type: none"> ▪ Support the preparation of the programme of activities of the PARC partnership, with particular focus on inventory of existing data and common knowledge on chemicals in order to identify gaps in data and anticipate future trends. ▪ Support to development of an Early Warning System for Europe ▪ Support to DG ENV / JRC for the review of the list of WFD Priority Substances, the Watch List for surface water and the Watch List for groundwater, Watch list for the Soil monitoring law. ▪ Support other Member States or River Basin authorities if requested. ▪ Links with WG-8: Support prioritisation of CECs as Sea Specific Contaminants in the marine environment (OSPAR, HELCOM, MED POL, Black Sea Commission). ▪ Links with WG-5 for identification of priority contaminants in treated wastewater intended for reuse ▪ Links with WG-7 on Soil and the Terrestrial environment for prioritisation activities. ▪ Link with the NTS CWG, in particular as regards DSFP, SLE and the definition of a mechanism for updating the NORMAN List of emerging substances as top priority substances of SusDat. ▪ NORMAN MassBank: prioritisation of relevant compounds whose mass spectra are not yet included in this online repository. ▪ Interlinking with the entire NORMAN Database System.
Participants	INERIS, EI, UBA, Ecotoxcentre, NKUA, LCSB, OVAM, UFZ, KWR, DERAC (All WG-1 are welcome)
Proposed	in-kind In-kind contribution is foreseen in all proposed activities as a complement to NORMAN funding.

contribution	<p>INERIS: coordination activities; contribution in dataset validation for modelling, conceptualisation of the prioritisation scheme, interface / visualisation; use categories for substances;</p> <p>UBA: coordination activities, with a focus on Ecotox and machine learning Expert Group; compilation of datasets and their validation for modelling, conceptualisation of the prioritisation scheme, interface / visualisation; DCT development, use categories for substances;</p> <p>EI: programming of the database modules, derivation of prediction models; running the predictions for SusDat compounds; running of suspect screening workflow (DSFP) for prioritisation case studies;</p> <p>NKUA: programming;</p> <p>LCSB: use categories for substances: preparation of meetings, communication with other data sources</p> <p>Ecotoxcentre: contribution in compilation of existing PNECs and validation of datasets for the ECOTOX module;</p> <p>DERAC: compilation of existing PNECs and connection with WG-5 for prioritisation of substances in matrices of interest for "reuse";</p> <p>UFZ: conceptualisation of the modules, interface / visualisation; use categories for substances;</p> <p>OVAM: conceptualisation of the prioritisation scheme for soil.</p>
Contribution needed from NORMAN Association¹	<p>Total net budget under JPA 2025: 25000€ (of which 9000€ postponed from 2024)</p> <p>Task 1: Support the prioritisation work of the Commission services at European level and comment on relevant documents and queries (PARC, EWS, where relevant). (in kind)</p> <p>Task 2: Collection and prediction of compound-specific information in support of prioritisation (UBA / EI / NKUA / DERAC / INERIS) (ongoing activity to be pursued in 2025)</p> <ul style="list-style-type: none"> - Retrieval and harmonisation of ecotox data from external databases (2000€) - Compilation of experimental data and model predictions of hazardous properties (2000€) <p>Task 3: Collection of existing PNECs and deep learning-based toxicity predictions for use in NORMAN PNECs (UBA / EcotoxCentre / DERAC / EI) (ongoing activity to be pursued in 2025)</p> <ul style="list-style-type: none"> - Model development for predictions of acute and chronic toxicity values for additional 10 species (2000€) - Application of the models for existing and new compounds in SusDat (2000€) - Compilation and uploading of regulatory PNECs (1000€) <p>Task 4: Consolidation of the Phys-Chem and Hazard properties module and integration in the NORMAN Database System and Prioritisation Tool (UBA, INERIS, EI in collaboration with WG-1 partners)</p> <ul style="list-style-type: none"> - Contribution from NORMAN is already included under the NDS Factsheet <p>Task 5: Compilation of data / information regarding 'Use categories' and 'Chemical Functional Use' for all SusDat compounds (LCSB / UBA / INERIS / EI)</p> <ul style="list-style-type: none"> - Retrieval and organisation of "Use" information from PubChem and other sources, programming, meetings (LCSB: 9,000€ for a 6 month project at €1,500/month for a student) <p>Task 6: Prioritisation framework: follow-up activities to integrate and test the new workflow as part of the Prioritisation Tool in the NDS (EI/ INERIS / UBA)</p> <ul style="list-style-type: none"> - Programming work of the back-end and interface of the extended prioritisation module (3000€) <p>Task 7: Improvement of the features of the prioritisation tool (expert consultation and programming activities) (ALL) (EI, INERIS, UBA, UFZ, NKUA, OVAM, DERAC, LCSB, KWR)</p> <ul style="list-style-type: none"> - Design and update of the interface for visualisation of the prioritisation results (3000€) <p>Task 8: Prioritisation of data from 5th Joint Danube Survey 5 (collaboration with ICPDR) (INERIS, UBA, EI in collaboration with WG-1 partners) (end of 2025 – 2026)</p> <ul style="list-style-type: none"> - Data processing, drafting report, meetings (in kind) <p>Task 9: Testing state of the environment indicators in the Danube River Basin (NORMAN collaboration with 5th Joint Danube Survey) (INERIS, UBA, EI and ALL) (2025-2026)</p> <ul style="list-style-type: none"> - Data processing, drafting report, meetings (in kind) <p>Task 10: NORMAN contribution to Early Warning System (UBA, INERIS, and partners willing to be involved)</p> <ul style="list-style-type: none"> - Development of the concept and design and sub-group meetings (1000€) <p>Task 11: Collaboration with other WGs to support prioritisation of contaminants in new compartments beyond the freshwater compartment and across compartments and for specific chemical groups (in kind) (budget included in the respecting WG factsheets)</p>

¹ 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.)