Preprint Title Instance maps as an organising concept for complex experimental workflows as demonstrated for (nano)material safety research

Authors Benjamin Punz, Maja Brajnik, Joh Dokler, Jaleesia D. Amos, Litty Johnson, Katie Reilly, Anastasios G. Papadiamantis, Amaia Green Etxabe, Lee Walker, Diego S. T. Martinez, Steffi Friedrichs, Klaus M. Weltring, Nazende Günday-Türeli, Claus Svendsen, Christine Ogilvie Hendren, Mark R. Wiesner, Martin Himly, Iseult Lynch and Thomas E. Exner

Publication Date 23 Apr. 2024

Article Type Full Research Paper

ORCID® iDs Benjamin Punz - https://orcid.org/0000-0001-9662-4739; Maja Brajnik - https://orcid.org/0000-0002-7420-1388; Joh Dokler - https://orcid.org/0000-0002-8053-8198; Jaleesia D. Amos - https://orcid.org/0000-0002-9769-4920; Katie Reilly - https://orcid.org/0000-0002-6054-0645; Anastasios G. Papadiamantis - https://orcid.org/0000-0002-1297-3104; Amaia Green Etxabe - https://orcid.org/0000-0003-4134-163X; Steffi Friedrichs - https://orcid.org/0000-0002-7276-892X; Nazende Günday-Türeli - https://orcid.org/0000-0002-6310-4785; Martin Himly - https://orcid.org/0000-0001-5416-085X; Iseult Lynch - https://orcid.org/0000-0003-4250-4584; Thomas E. Exner - https://orcid.org/0000-0002-1849-5246

Instance maps as an organising concept for complex experimental workflows as demonstrated for (nano)material safety research

Benjamin Punz[1], Maja Brajnik[2], Joh Dokler[2], Jaleesia D. Amos[3], Litty Johnson[1], Katie Reilly[4], Anastasios G. Papadiamantis[4], Amaia Green Etxabe,[5] Lee Walker,[5] Diego S.T. Martinez[6], Steffi Friedrichs[7], Klaus M. Weltring[8], Nazende Günday-Türeli[9], Claus Svendsen[5], Christine Ogilvie Hendren[10], Mark R. Wiesner[3], Martin Himly[1*], Iseult Lynch[4*], Thomas E. Exner[2,11*]

[1] Department of Biosciences & Medical Biology, Paris Lodron University of Salzburg, Hellbrunnerstrasse 34, 5020 Salzburg, Austria;

[2] Seven Past Nine d.o.o., Hribljane 10, 1380 Cerknica, Slovenia;

[3] Center for the Environmental Implications of Nano Technology (CEINT), Civil & Environmental Engineering, Duke University, Durham, North Carolina, 2770y8, USA;

[4] School of Geography, Earth and Environmental Sciences, University of Birmingham, Edgbaston, B15 2TT Birmingham, United Kingdom;

[5] UK Centre for Ecology and Hydrology, Pollution, Wallingford, Oxfordshire, United Kingdom;

[6] Brazilian Nanotechnology National Laboratory (LNNano), Brazilian Center for Research in Energy and Materials (CNPEM), Campinas, Sao Paulo, Brazil;

[7] AcumenIST SRL, Etterbeek, Belgium;

[8] Gesellschaft für Bioanalytik Münster, Mendelstraße 17, 48149 Münster, Germany;

[9] MyBiotech GmbH, Industriestrasse 1B, 66802 Überherrn, Germany;

[10] Department of Geological and Environmental Sciences, Appalachian State University, Boone, USA;

[11] Seven Past Nine GmbH, Rebacker 68, 79650 Schopfheim, Germany.

* Corresponding authors: Martin Himly – Martin.Himly@plus.ac.at; Iseult Lynch – i.lynch@bham.ac.uk; and Thomas Exner – thomas.exner@sevenpastnine.com

Abstract

Nanosafety assessment, which seeks to evaluate the risks from exposure to nanoscale materials, spans materials synthesis and characterisation, exposure science, toxicology, and computational approaches, resulting in complex experimental workflows and diverse data types. Managing the data flows, with a focus on provenance (who generated the data, for what purpose) and quality (how the data was generated, using which protocol, with which controls), as part of good research output management, is necessary to maximise the re-use potential and value of the data. The concept of Instance Maps has been developed and evolved to visualise experimental nanosafety workflows and to bridge the gap between the theoretical principles of FAIR (Findable, Accessible, Interoperable and Re-usable) data and the everyday practice of experimental researchers. While Instance Maps can be created for completed studies or publications, they are most effective when applied at the study design stage to associate the workflow with the nanomaterials, environmental conditions, method descriptions, protocols, biological and computational models to be used and the data flows arising from study execution. Application of the Instance Map Tool (described herein) to research workflows of increasing complexity is presented to demonstrate its

utility, starting from assessment of nanomaterials transformations in complex media, and documentation of nanomaterial synthesis, functionalisation and characterisation, to description of the culturing of ecotoxicity model organisms *Daphnia magna* and their use in acute and chronic standardised tests for nanomaterials ecotoxicity assessment, and visualisation of complex workflows in human immunotoxicity assessment using cell lines and primary cellular models. These examples showcase the use of the Instance Map approach for coordination of materials and data flows in complex multi-partner collaborative projects, and to support demonstration case studies. Finally, areas for future development of the Instance Map approach and tool are highlighted.

Keywords

Study design; experimental workflow visualization; data collection and quality control; FAIR; nanomaterial life cycle stages; data provenance

Introduction

The manipulation of matter at the nanoscale, and the emergence of nanoscale materials whose properties can be tailored by changing their size, shape, surface chemistry, and functionality has led to the designation of nanomaterials as a key enabling technology and to their subsequent inclusion in the broader categorisation of Advanced Materials [1,2]. Applications of nanomaterials derive in many cases from their high surface reactivity, which results from their small size and large surface area, including applications in catalysis [3,4] (e.g., as catalytic converters in engines, for energy capture and storage), and as sensors [5,6] (e.g., for bioremediation, environmental monitoring). In medicine [7,8] and agriculture [9,10], their loading with active ingredients and targeting to key sites for action are enabled through surface

functionalisation and the small size of nanomaterials which allows them to access all areas. An important consequence of nanomaterials reactive surface area is their instantaneous interaction with their surroundings, through formation of an acquired environmental or biomolecule corona [11,12] and/or via physical or chemical transformations that can occur at any of the nanomaterials life cycle stages [13,14].

The ability of engineered nanomaterials to change characteristics based on the properties of their environment presents a unique challenge for evaluating their potential environmental and human risks [15,16]. This “context dependence” of many nanomaterials properties requires distinction between *extrinsic nanomaterial properties* that can change as the surroundings change (such as zeta potential which depends on the pH and ionic strength of the surrounding medium [17]) and *intrinsic nanomaterial properties* which are not affected by the surroundings, such as bandgap and structural arrangement [18]. This tendency of nanomaterials to change with their surroundings, or even with time during storage [19], suggests that the time between synthesis and initial characterization and toxicity or analysis, as well as changes in conditions of the surrounding medium, are important to document, although is not routinely reported in the literature [20]. Bear et al. suggested that the essential history of a set of particles can be identified as *provenance information* that tells the origin or source of a batch of nano-objects along with information related to handling and any changes that may have taken place since it was originated [21], which would play a useful role in decreasing the extent of particle variability and the lack of reproducibility observed by many researchers.

Efforts to capture and document batch-to-batch variability of nanomaterials synthesis routes were made in the QualityNano project [22], and to establish a Uniform Description System for nanomaterials to describe nanomaterials (batches) uniquely, so that each is differentiated, and to determine when two (batches of) nanomaterials

are equivalent to whatever degree desired [23]. Given the fact that Nanomaterials similarity can only be verified through extensive physicochemical characterization, which is often done in parallel to toxicity testing, a work-around solution was proposed whereby projects could assign a unique identifier to their batches of nanomaterials via the European Registry of Nanomaterials [24] and add the characterization data later, thus enabling batch similarity to be assessed by users wishing to integrate datasets. However, it is not clear that characterization data is added in practice, or whether any of the approaches suggested to date have been applied in a practical sense by the nanosafety research community. This could in part be due to the breadth of the nanosafety research domain, whereby it is often not the researchers that produced or characterized the nanomaterials that are undertaking the different steps of the exposure or hazard assessment. Indeed, this effect of specialization was observed in reporting of studies of nanomaterials protein coronas, whereby the documentation of the nanomaterials dispersion and corona formation steps was very complete, but the description of the protein isolation and informatics steps was much less complete. This gap in documentation was attributed to the fact that the omics analyses are often performed by core facilities and nanomaterials researchers don't know exactly what needs to be documented about these steps to enable the study to be reproduced [25]. Another frequently encountered challenge is the misconception that a statement regarding use of a standard test guideline or guidance document is sufficient as metadata about a nanomaterials toxicity study to enable re-use of the resulting data. Notably these standard tests, as developed by the Organisation for Economic Cooperation and Development (OECD) are usually quite broad, as they are globally agreed, and thus allow users some flexibility in terms of medium, soil, dispersion approach and so forth, meaning that detailed documentation of each step is still required to allow others to re-use the data with confidence. This is especially important

for nanomaterials, given that the test guidelines originally developed for soluble chemicals are currently being revised for use with nanomaterials [26].

The complexity and transformability of nanomaterials also has consequences for the databases used to organise and store nanomaterial characterisation and (eco)toxicity data: databases needed to adapt to the nature of the data they were required to store. One innovative approach, taken by the NanoInformatics Knowledge Commons (NIKC) database [27], was to introduce the concept of the "*nanomaterial instance*" to capture the transformations that nano-scaled materials undergo in environmental and biological compartments as a visual representation to guide the data curation process [28], i.e., to highlight where changes to the nanomaterial may have occurred and thus where additional characterisation information would be needed. Instances were designed to capture the necessary metadata needed to describe a material and its surrounding medium in mesocosm experiments, while keeping the sequence of transformations intact (e.g., a material deposited in soil resulting in the material's uptake by surrounding plants, which are then eaten by insects). Material transformations are tracked through connected instances. As originally conceived, the nanomaterial instances were used to systematically retrofit experimental data from published literature describing nanomaterials mesocosm studies in order to capture the nanomaterial transformations in a manner that sufficiently includes surrounding medium characteristics, thus representing both intrinsic and extrinsic properties of the studied material [20]. Mesocosm studies are often complexly layered with multiple assays and characterisation methods occurring sequentially or concurrently, often within a larger encompassing study in order to gain a more complete understanding of nanomaterial behaviour. The NIKC curation team was tasked with translating these experimental studies into nanomaterial instances and identifying important metadata associated with each instance. This was done by categorising experimental data into

one of five categories: an instance, material, medium, property, or supplementary; where a property can describe either a medium (e.g., environmental, biological, experimental) or material, “supplementary” provides a way to include visual information about a property (e.g., image, diagram), and the instance itself is the point in time when the material, medium, and properties are being described together. A study could have as many instances as needed to describe each of the potential material transformations. For quality assurance and quality control (QA/QC) purposes, the curation team needed a way to compare defined instances and transformations. After many trials, the most efficient method for curation was a visualisation or map that the curators would follow during the curation process; thus, instance mapping was created. More information on the approach is available in Amos *et al*. [28].

The benefits of such a visual representation for study design to guide researchers in what characterisation and system metadata was needed for complete reporting of nanosafety studies emerged quickly, with researchers using instance maps independently of the NIKC for purposes beyond data curation. As a project planning extensive mesocosm studies, NanoFASE adopted the concept for their mesocosm study reporting. In collaboration with their NanoCommons data shepherd [29], The NanoFASE project adopted the instance map approach for project-wide data management, to structure the data reporting of the complex mesocosm experiments using a modified version of the NIKC file format and uploading the data onto the NanoCommons Knowledge Base [30]. The early instance maps were drawn by hand, without tools specifically designed to create these maps, and thus their use as an integral part of the overall data management infrastructure emerged holistically and bottom-up and evolved based on real application by the nanosafety research community.

Much of the potential benefit provided by the use of instance maps arises from removing the current separation of data production from data curation, harmonisation, reporting and FAIRification (making data Findable, Accessible, Interoperable and Re-useable). Instance maps represent an integral part of data production following an on-the-fly data management approach [31], supporting all stages of the data management life cycle [29] by allowing the easy creation of a visual draft of the experimental workflow at the study design phase and then associating this workflow with the materials, environmental conditions, method descriptions, protocols, biological and computational models to be used and the data produced during study execution. Indeed, this use of instance maps to inform the earliest parts of the data lifecycle was a primary goal of the NIKC team in developing the approach, in order to generate "premeditated interoperability" of resulting datasets and therefore enable broad integration of datasets across multiple groups; however, the realisation of that goal could only emerge upon adoption of the approach by other research groups. The NanoCommons project pioneered use of instance maps for documenting study design and data capture needs as part of the data shepherding approach, and developed a software tool for the creation of instance maps. The approach has now been taken up and continued in MACRAMÉ and other recently funded advanced materials projects. As demonstrated here, use of instance maps to visualise material transformations has evolved into a powerful tool that extends beyond curation, and beyond engineered nano-scaled materials. Indeed, researchers have started using instance maps to aid the design and planning of experiments, as communication and instructional tools at individual and collaborative levels, and in educational settings.

This paper presents examples of such new applications of instance maps for planning, documenting, and sharing study designs and associated data and metadata. The instance map tool allows the user to design workflows in a fully customised manner

and to connect the nodes (instances, properties, protocols, and data) with protocols and data management tools such as electronic laboratory notebooks (ELNs), which should aid interoperability. While the focus of the cases presented here is nanosafety and sustainability, the general utility and applicability of the instance map concept to describe complex experimental and computational studies in other research areas and potentially in regulatory settings and industry-applied development and innovation processes is also evident.

Methodological approach

Definition of the instance map concept

The original instance maps, used as organisational structure in the data curation efforts for the NIKC database [27], enabled users to visually document nanomaterial transformations while capturing the necessary metadata [28]. The experimental data is sorted into five categories: instance, material, medium, property, and supplementary, to catalogue the metadata describing the nanomaterial and the exposure medium. An instance is defined as the nanomaterial in a medium at a specific moment in time. The material and medium categories are used to describe the instance. A physical or chemical change to the nanomaterial that (potentially) alters the physico-chemical or biological properties of the material results in a new instance. An instance map then represents a flow chart of the nanomaterial fate represented as a directed, often tree-like graph built out of nodes connected by edges, represented as arrows to show the directionality [28]. The main branch(es) is(are) formed by consecutive instances and other branches connected to the main branch describe the material and medium at this specific point in the experiment and their properties (see Figure 1).

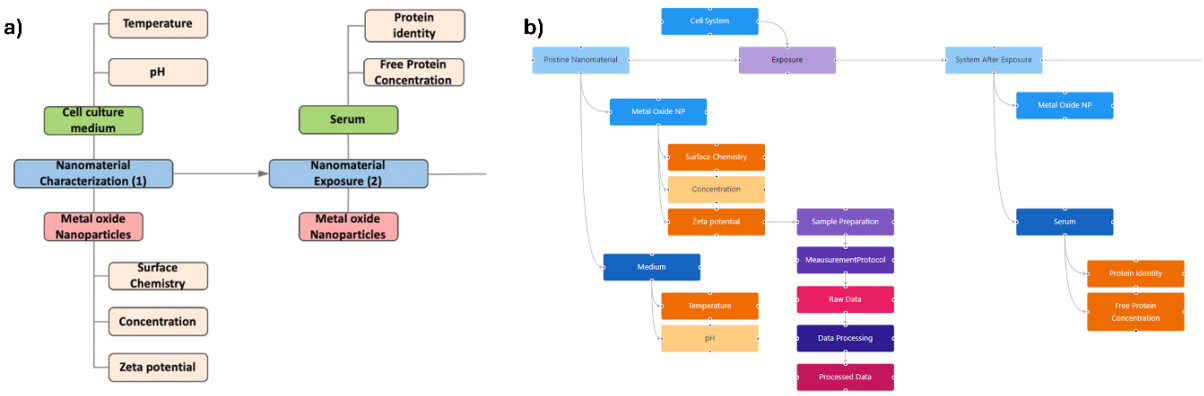

Figure 1: Comparison between a) the original conception of an instance map using the original definition from NIKC, modified from Amos et al. [27], and b) an instance map generated using the Instance Map tool with its extended node library. The full instance map in b) is accessible at https://instance-maps.stage.sevenpastnine.com (username: SupportingInfo, password: maps-for-paper) for interactive inspection.

In its original conception, the chosen categories (also called nodes), and the strict set of rules on how to place and connect the nodes was optimised for the needs of the NIKC data curators, and later for describing the mesocosm experiments of the NanoFASE project and the corresponding data curation template. The NanoCommons data shepherding services facilitated other research groups to reuse instance maps to describe their research [33, 33]. These reuses also showed that a few extensions and the provision of a specialised software tool to create the maps would further facilitate and encourage adoption for other types of experiments, new use cases and application of instance maps as a tool to optimise and document study design.

The first extension proposed to support study design was to differentiate between different types of properties (see Figures 1 and 2). In the NIKC curation efforts, all data were extracted from scientific publications and, thus, there was no obvious separation in the eyes of the curator between data produced specifically within a paper (primary data) or data taken from literature or public databases (secondary

data). This distinction becomes important, however, when using instance maps for complex study design workflows, where primary data can be further categorised into wet-lab and computationally produced data. To capture the complete experimental metadata, it was also seen as beneficial to be able to explicitly refer to protocols (exposure, characterisation or toxicity), since going from one instance to the next can be a multi-step process involving application of numerous protocols and/or standard operating procedures (SOPs) of different origin. While this could be achieved by adding an instance for the resulting material state after each sub-step, these intermediate instances are not typically characterised experimentally and thus, the instances would make the maps more complex without adding much information. Explicit protocol nodes, in contrast, can be linked to the corresponding resources documenting the steps in the form of text documents, protocol repository entries or ELN pages. For data produced in the study, experience from the ACEnano project was exploited, which conceptualised a strong linkage between protocols and data using a workflow with stages for sample preparation, measurement, raw data (collection), data processing and processed data. Putting this all together, the Instance Map tool supports 12 nodes, grouped into 4 categories, as shown in Figure 2.

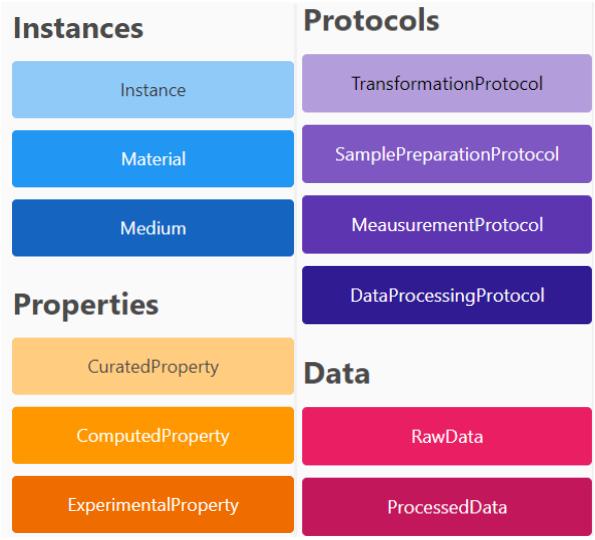

Figure 2: The nodes available in the instance map tool to represent a study, grouped into 4 categories. The Instance consists of the material and its medium (surroundings). Properties can be curated (from literature) or calculated (computed) or experimentally determined. Protocols cover all steps of the workflow, including any transformations, sample dispersion and exposure, measurement steps (e.g., physicochemical characterisation, (eco)toxicity evaluation, functional testing) and data processing such as gap-filling, data cleaning, statistical analyses etc. Data is then classified as raw (coming directly from the measurement) or processed (following steps such as subtraction of medium blanks or calculation of half maximal effect concentrations).

The NanoCommons instance map tool

A first prototype of an instance map service was developed, which speeds up the creation of the maps and allows linking of nodes to protocols and data sources. The tool is located at https://instance-maps.stage.sevenpastnine.com and can be accessed with a username and password (the maps of this publication can be accessed under username: SupportingInfo and password: maps-for-paper). The following functionalities are available:

1. Creation and modification of instance maps and provision of basic metadata.

2. Linking of data and other research outputs to individual nodes.
3. Sharing of instance maps with other users in the same user group, who can view the map and all associated data and (meta)data by accessing the Instance Map tool but cannot modify the maps.

The instance map tool was developed using a set of open-source frameworks and libraries. At the heart of the tool is ReactFlow for building node-based editors and interactive diagrams. ReactFlow is incorporated into the tool using the F#/Fable toolkit Feliz. For the backend, the Django framework is used alongside the relational PostgreSQL database to handle data storage and user management. During the NanoCommons project, a group of test users were engaged in assessing the tool's usefulness and interface usability. Regular feedback during all phases of the development was crucial in guiding the development process with regards to defining and prioritising the requirements in terms of nodes and edges.

An instance map can be created with a simple drag and drop action of the items (nodes). Users can choose between the twelve different types of nodes described above, which are grouped and colour coded for easier interpretation of map overviews. Individual nodes can be connected with edges to represent complete workflows. Data support is still limited in the current version of the tool but will be improved in the future to support the harmonised and interoperable on-the-fly data management concept envisioned in the introduction and described in [3]. Users can provide further information like descriptions, keywords, version numbers, creation dates, licences, contributors and references for the complete map as well as for individual nodes, as well as links (URLs or relative paths) to data files. This approach was chosen in the test phase to allow users greater flexibility with respect to the format in which their data is stored. Currently used formats for protocols and data include data serialisation formats like JSON and YAML, notebook pages (e.g., electronic lab notebooks like

SciNote, Jupyter and Colab computation notebooks), text documents (Microsoft Word, Google Docs), spreadsheets (Microsoft Excel) and provider-specific data files. Other possibilities include images, videos or links to public repositories. A demonstration of the tool is available at https://nanocommons.github.io/user-handbook/data-management/instance-maps/ along with a tutorial to support users.

Results and Discussion

The utility of the instance map service is demonstrated on a range of experimental workflows applied in nanosafety and sustainability assessment, representing assessment of nanomaterials or advanced materials via different endpoints and workflows. Typically, the overall experimental workflow in nanosafety assessment consists of (but is not limited to) some or all of the following steps: (i) material synthesis or procurement; (ii) further modifications, e.g., surface functionalization; (iii) a plethora of characterisation steps by physicochemical methods, potentially also including the application of computational modelling and prediction tools; (iv) determination of diverse biological endpoints *in vitro* and/or *in vivo*, which can also consist of both experimental and computational approaches, and (v) processing of the raw data and enrichment of the processed data and its integration to support risk assessment and/or safe by design applications.

Documenting nanomaterials synthesis and provision of unique identifiers for nanomaterials

Instance maps were used as a tool to visualise the synthesis of different types of surface modified nanomaterials. These maps were used to highlight how slight changes in the synthesis process can alter defining characteristics of the particle,

which may drastically change particle behaviour in environmental and biological media and nanomaterial (eco)toxicity. Although multiple instance maps were created for different types of surface-modified nanomaterials, only one is presented here (Figure 3). The synthesis method illustrated was published by Levard et al. [34] and was chosen because of its thoroughly described synthesis protocol and characterisation methods. It was also chosen because the nanomaterials were later used in an extensive exposure study examining toxicity responses of organisms based on differences in the particles' sulfidation levels [35]. We note that the same reasons underpinned its selection for discussion of instance maps in Amos et al. (2024) [28].

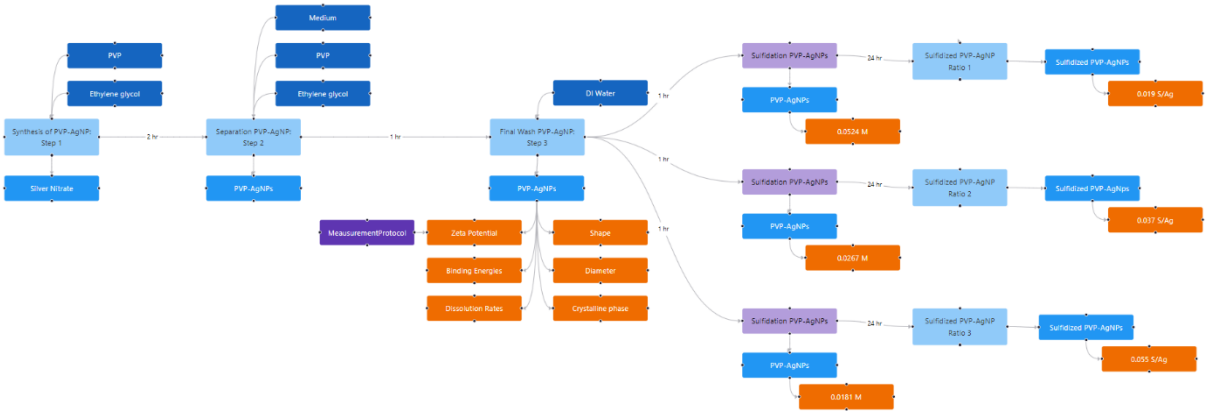

Figure 3: This instance map shows the synthesis protocol for the sulfidation of AgNPs, which was originally published in Levard et al. [34]. The map can be described in two phases. The first phase is the synthesis of silver nanoparticles (AgNPs) functionalized with polyvinylpyrrolidone (PVP), which occurs in the first three instances (shown as light blue nodes). Once the NPs have been synthesised, physical and chemical characterisation of the particles is performed. These characterisation endpoints can be seen in orange linked to the material in the third instance. The second phase is the sulfidation process, which can be seen in the light purple boxes. Although there are three 'transformation protocols' listed, the protocol is the same except for the concentration of PVP-AgNPs used. The map shows that as the concentration of PVP-AgNPs used in the sulfidation process decreases, the S/Ag ratio increases. The full

instance map is accessible at https://instance-maps.stage.sevenpastnine.com for interactive inspection (username: SupportingInfo and password: maps-for-paper).

The instance map in Figure 3 delineates all steps of the synthesis of sulphatised silver nanoparticles (AgNPs). AgNPs are synthesised with a polyvinylpyrrolidone (PVP) surface, by reduction of silver nitrate in ethylene glycol with 10k PVP. The PVP-AgNPs are characterised for some of their physical attributes such as the particles' shape, size, and crystalline phase. The particles are then sulphatised using different specified concentrations of the PVP-AgNPs resulting in increasing levels of sulfidation measured by the S/Ag ratio. Thus, four different NPs need to be distinguished, and tracked in the subsequent toxicity experiments, leading to a need for unique identifiers for the nanomaterials.

Instance mapping is thus being extended to support and implement emerging standards to FAIRrify nanomaterials data by creating a common naming convention. An international and interdisciplinary group is currently working on refining a standard nomenclature for nanomaterials, the InChI for nano or NInChI [36], based on the International Chemical Identifier (InChI). The objective is to create a notation that is readable to both humans and machines that encompasses chemical and physical attributes of the material. As shown in the example in Figure 3, nanomaterials are often layered, often with a core and a functionalized surface which can be engineered for specific purposes and can modulate toxicity endpoints. Ideally, a nomenclature would include details on chemical identities of the nano core, surface and its transformation where it is the transformed form of the nanomaterial that is being evaluated, any impurities, physical descriptors of the material's morphology, as well as the nature of the bonds between the surface and core. This level of detail can only be gained by

understanding how the nanomaterial is synthesised, which is where instance maps will be a critical tool.

Monitoring nanomaterial transformation in complex environmental media

Ecotoxicity exposures conducted in soil and mesocosm experiments are often complex with multiple parameters and endpoints (e.g., [37–40]). The diversity of data types required to monitor soil, porewater, nanomaterials and organisms requires many sample collections and analyses, as well as ensuring collection of pre-experiment data and metadata. The complexity of the experiments is simplified by the use of instance maps, which allow an overview of biological and chemical sampling during the mesocosm experiment. By detailing all relevant metadata and post exposure analyses, instance maps visualise the flow of data collection and methodologies, including the biological culture information and chemical pre- and post-exposure data (Figure 4A).

The instances in this example (see Figure 4) follow the timeline of exposure, and at each instance the nodes depict the data pertaining to that particular instance. Instances at the top of the map (see Figure 4 ‘Set up – nanomaterial dispersion and soil spiking’) occur before the exposure to organisms, and include the nanomaterial dispersion and their addition to soil, followed by instances detailing the addition of organisms and then the timepoints of sample collection. Data was organised left to right to visualise the distinction between curated data and/or any pre-experiment information and data generated by the experiment itself (Figure 4B). The data and processes are visualised as nodes attached to each instance. On the left side of this example are nodes relating to the components prior to their addition to the experiment. In this case, the information on the pristine nanomaterial, medium and soil, includes

suppliers, batch numbers, CAS numbers, pH and any other pre-processing steps before their addition into the experiment (mesocosm). For the first exposure instance, also the species information such as the cultivars, suppliers, culture maintenance information, and QC of organisms entering the experiment are included.

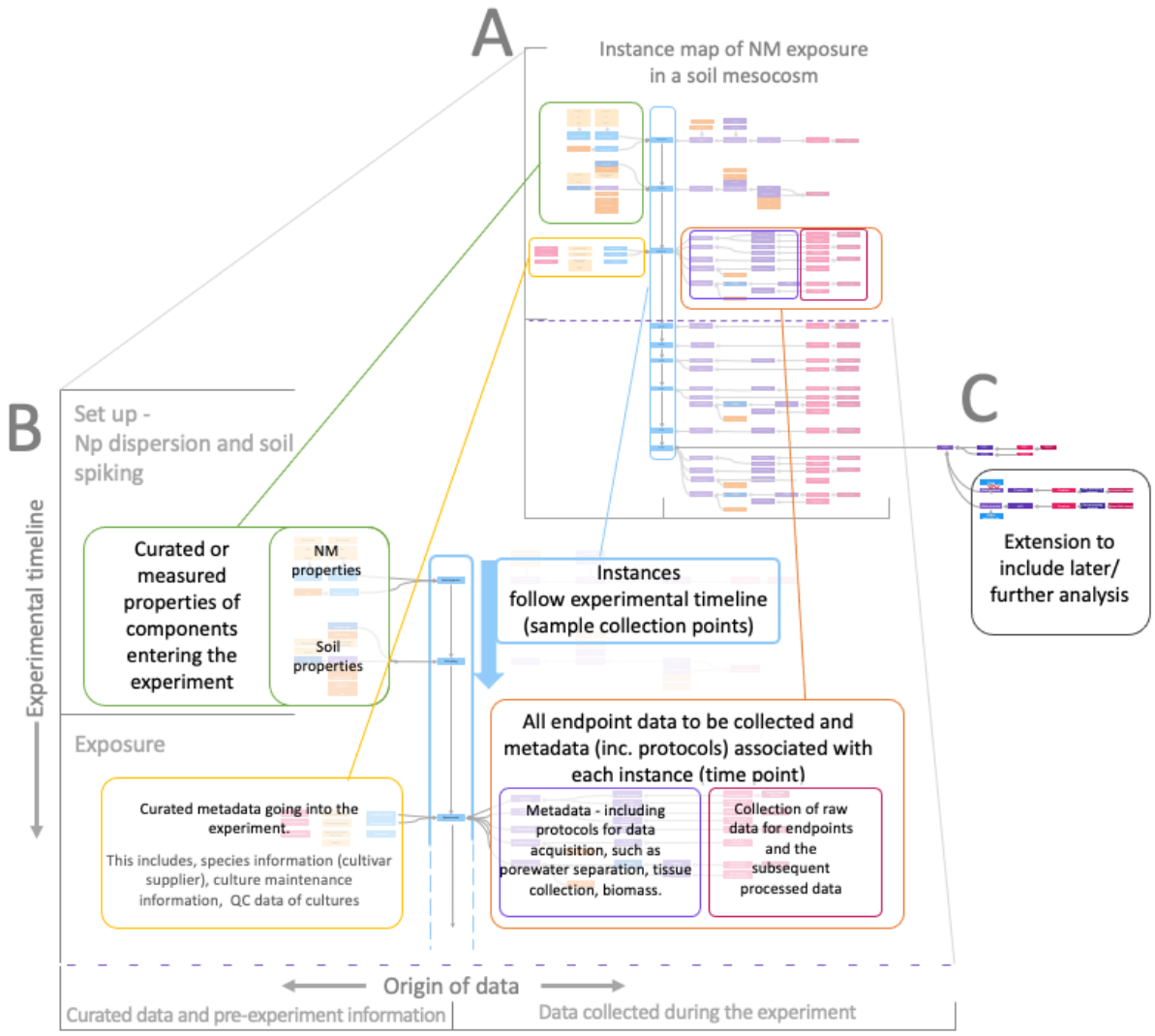

Figure 4: Instance map for a nanomaterials mesocosm experiment. A) Representation of an instance map for a mesocosm exposure experiment. B) An expanded map region to visualise the experiment organisation and the flow of data collection. Instances (blue boxes outlined with a blue border) are organised in time from the top of the map (blue arrow represents direction of time). Data is split on either side of the instances to

distinguish its origin. To the left are the green and yellow boxes that show curated data and pre-experimental information. Curated data and pre-experiment information is further split across instances to show when it is applicable to pre- (green) or post- (yellow) organism exposure. To the right side of the instances is an orange box that shows all data generated from a given instance. This data is also further split into two categories. Firstly, raw data and processed data (pink border) and, secondly, the methodologies and processing approaches used to derive that data (purple border). C) Extension of a sample node to include further analysis and data points. The full instance map is accessible at https://instance-maps.stage.sevenpastnine.com for interactive inspection (username: SupportingInfo and password: maps-for-paper).

On the right of the close-up (Figure 4B), an organised display of any data generated by the experiment itself is shown. This is split into raw and processed data, as well as the processes of their collection. The pink section nodes represent the raw data sets collected, such as the pH of soil after the addition of the nanomaterial, organism biomass data, metal concentrations in organism tissues, and any processed data which derives from this raw data, such as EC50s or metal bioaccumulation rates into organisms. Information regarding the protocols and methods used for data collection and how samples were processed is also attached to the data. For example, soil porewater separation protocols needed to help generate porewater metal concentration data, and all tissue sample collection processes are available.

The level of overview provided by instance maps greatly benefits the complex, multi-endpoint, experiments common to ecotoxicology, ensuring metadata collection, optimal experimental design, informing sample processing schedules and data management plans. The flexibility of the instance map system means maps can be extended to include any further branching to processes, such as adding any later

analysis of collected samples by extending a branch for that sample, e.g. transcriptomic analysis on exposed organisms by RT qPCR (Figure 4C).

Linking assay QA/QC with SOPs for running cultures of biological organisms and standardised ecotoxicity testing

Keeping records of the normal organism behaviour in individual labs is vital for regulatory testing but is not something that is formalised in most academic laboratories. Thus, instance maps can also be used to build awareness of the pre-experiment steps and the importance of documenting these as they form part of the provenance and QA/QC metadata that underpins regulatory testing. This data supports demonstration of the trustworthiness of (hazard) data to others who may wish to re-use the data, for example in modelling or as part of a risk assessment.

The model organism *Daphnia magna* is cultured in a high hardness medium, which is aerated for a minimum of 8 hours prior to use in culturing and the dissolved oxygen content is measured every 2-3 days to ensure it stays within the acceptable range. The pH of the medium is also measured and moderated to within the defined parameters for the specific medium before use for the ongoing culturing of daphnia. The running cultures are typically in large (1L) beakers with 900 mL medium and can contain 10-15 adults, with the medium being refreshed three times per week. All cultures are fed the same daily algal ration of *Chlorella vulgaris* (7.5 mg C days 0-7, 11.25 mg C days 7 onwards, with double rations on Fridays to cover the weekend) and are kept in a 20 °C laboratory under a 16:8 hour light:dark cycle. The steps involved in maintaining the daphnia, and the algae on which they feed, are shown in the instance map of Figure 5. Third brood daphniids are used for all ecotoxicity experiments (e.g., acute or chronic toxicity testing) to ensure optimum genetic health of future cultures.

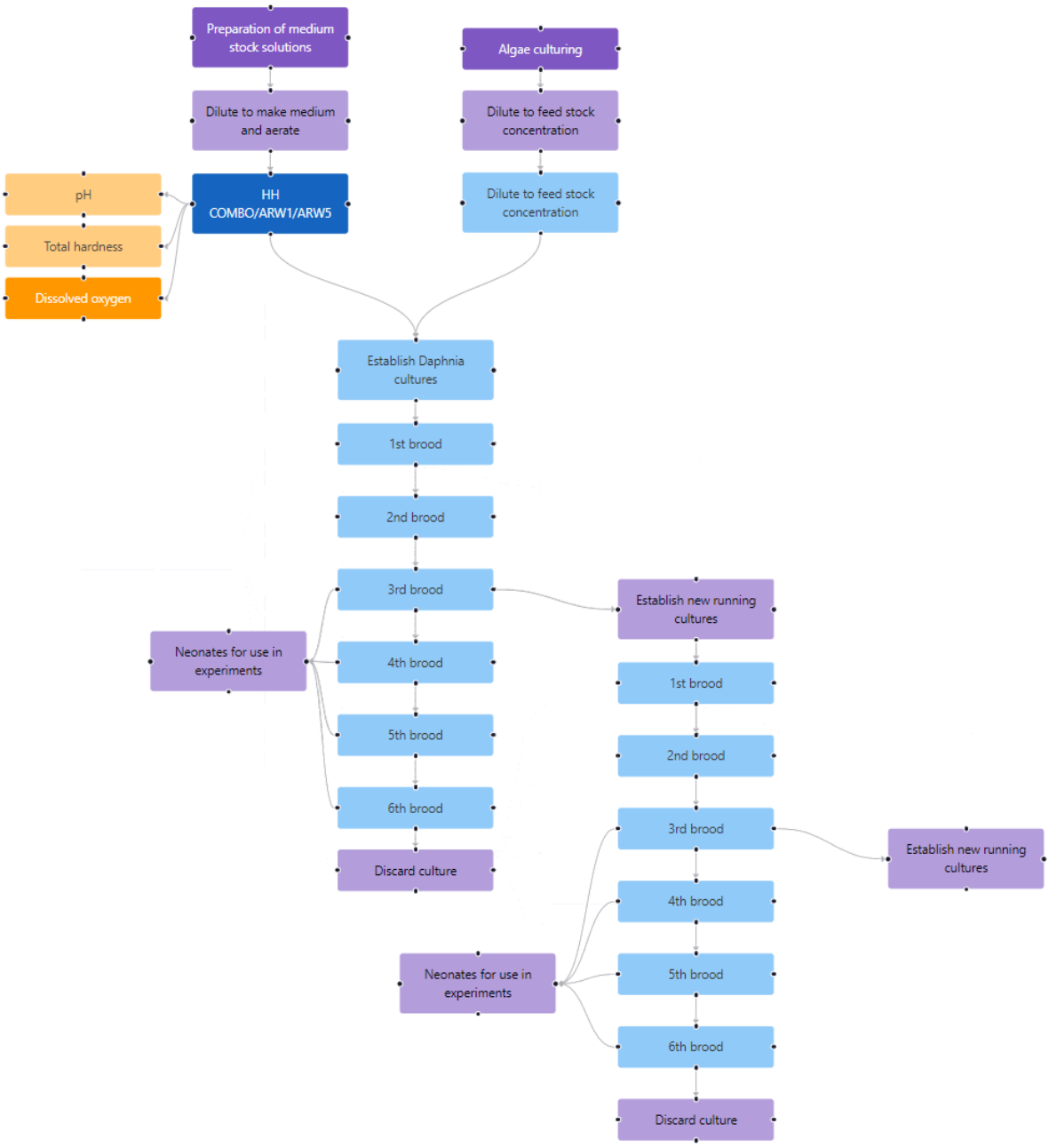

Figure 5: Instance map visualising the steps in maintaining continuous *D. magna* cultures. Daphnia typically produce broods from about 10 days old and roughly every 3 days thereafter, with the 3rd to 7th broods being the most genetically stable, and thus being suitable for ecotoxicity experiments. Tracking of the number of offspring per brood is one of the essential QC measures to record, using the template shown in Figure 6. Details such as organism species, strain, and culturing conditions (temperature, pH, dissolved oxygen, light:dark cycle) can be captured here as well as

the specifics of the medium, culturing vessels etc. The full instance map is accessible at https://instance-maps.stage.sevenpastnine.com for interactive inspection (username: SupportingInfo and password: maps-for-paper).

Date	Day of culture	Culture (number in jar)						Offspring						Food (mL)	Media change	Observation/comments
		1	2	3	4	5	6	1	2	3	4	5	6			
02/01/2024	1	12	12	12	12	12	12							0.75	✓	
03/01/2024	2	12	12	12	12	12	12							0.5		
04/01/2024	3	12	12	12	12	12	12							0.75		
05/01/2024	4	12	11	12	12	12	12							1		
06/01/2024	5	12	11	12	12	12	12							1	✓	
07/01/2024	6	12	11	12	12	12	12							2		
08/01/2024	7	12	11	12	12	12	12							0		
09/01/2024	8	12	11	12	12	12	12							1.5	✓	
10/01/2024	9	12	11	12	12	12	12							1.5		
11/01/2024	10	12	11	12	12	12	12							1.5		Eggs in brood pouch
12/01/2024	11	12	11	12	12	11	12	✓	✓	✓		✓	✓	1.5	✓	Start of first brood
13/01/2024	12	12	11	12	12	11	12				✓			3		
14/01/2024	13	12	11	12	12	11	12							0		
15/01/2024	14	12	11	12	12	11	12	✓	✓	✓		✓	✓	1.5	✓	
16/01/2024	15	12	11	12	12	11	12				✓			1.5		
17/01/2024	16	12	11	12	12	11	12							1.5		

Figure 6: A simple data capture template for monitoring the health and performance of running daphnia cultures, wherein the amount of food and dates of medium changes are reported along with the numbers of offspring measured per running culture jar. Long term tracking of culture performance allows confidence in data generated in regulatory testing using standardised assays, as any deviations from normal behaviour can be confirmed as being from the exposure rather than from any anomalies in how the test was performed.

The steps in the acute daphnia toxicity test, performed according to the OECD standard test guideline (OECD 202 [41]) have been visualised using an instance map, as per Figure 7. An intentional feature of the OECD test guidelines is that they leave some flexibility for the user, in that they recommend a specific medium, but it is not essential (and indeed many labs use tap water or bore hole water), and thus the lab needs to prepare its own detailed SOP that underpins the experiment. In the example shown, we have not linked to other aspects of an overall study that would be required, such as characterisation of the stock solution and assessment of the nanomaterials'

stability in the test medium. However, the beauty of the instance map approach is that this linking of experiments / experimental steps is easy. A related data capture template has been developed and is linked to the raw data node. The OECD 211 chronic (reproductive assay) has also been mapped, as shown in Figure 8, noting that the concentration used in the chronic test is usually derived from the acute test (e.g., the EC30 or EC10 concentration), and thus these instance maps can also be linked, and indeed linked to the running culture instance map of Figure 5.

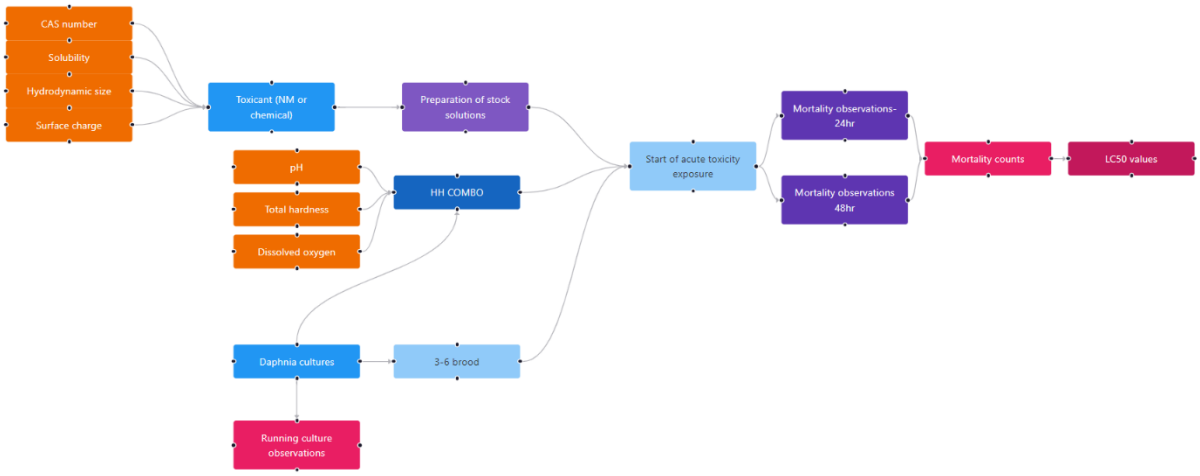

Figure 7: Representation of the OECD 202 Test Guideline for acute toxicity to daphnia as an instance map. For nanomaterials there would be an additional link from the stock solution to the range of characterisation studies needed, such as size distribution, surface charge, stability over time, etc. The full instance map is accessible at https://instance-maps.stage.sevenpastnine.com for interactive inspection (username: SupportingInfo and password: maps-for-paper).

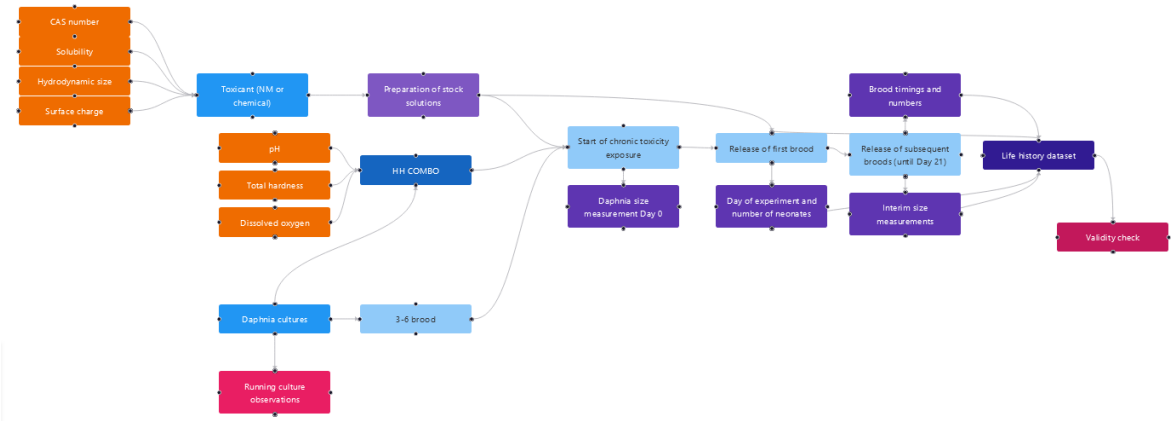

Figure 8: Representation of the OECD 211 Test Guideline for reproductive (chronic) toxicity to daphnia as an instance map. The exposure concentration is determined from the acute dose-response curve, generated according to the instance map shown in Figure 7, which in a next interaction of the Instance Map tool will be fully integrated. The instance map is accessible at https://instance-maps.stage.sevenpastnine.com for interactive inspection.

In line with the QA/QC efforts presented here, initiatives are ongoing at the European and to certain extent even global level aiming, *e.g.,* at harmonisation of nanomaterials characterisation reporting, its terminology, classification, and metadata. A standard structure containing this type of information relating to (i) materials characterization (meta)data, termed CHADA (CHAracterisation DAta and description of a characterisation experiment), has recently been proposed [42]. Standardised or harmonized reporting formats had previously been called for, such as listing of minimal reporting standards for biological assays studying the interactions of nanomaterials with biological materials, termed MIRIBEL [43]. The prime intention here is to improve future exchange of datasets among materials characterization experts, facilitate collaboration with industry end-users, and to optimize the interoperability of data and thus enable better data re-use by modelling experts. Likewise, efforts are ongoing for

harmonizing the (ii) materials MOdelling DAta terminology, resulting in templates for physics-based model description, termed MODA [44], driven by the activities of the European Materials Modelling Council (EMMC), resulting in a Workshop Agreement of the European Committee for Standardization (CEN). Instance maps can support this effort by graphically resolving reporting documents as they enable a structural representation of the experimental (or even computational) data workflow. In the context of biological experimentation, we can link the instance maps to (iii) biological data reporting that fulfils criteria such as advocated by MIRIBEL. Analogous to the two afore-mentioned reporting formats, such a biological documentation could be termed BIODA, a reporting structure for BIOlogical assay DAta. Such concepts will prove useful when highly complex workflows are built and ultimately data have to be aggregated from the different criteria. For example, regulatory readiness of testing pipelines based on new approach methodologies where batteries of >20 assays with >50 individual endpoints are compared and data (from different laboratories) needs to be aggregated [45]. The offspring tracking presented in Figure 5 may represent the first step towards implementation of a BIODA, to allow benchmarking and interoperability of data from different labs, similar to the CHADA and MODA concepts.

Showcasing complex workflows in human immunotoxicity assessment using cell lines and primary cellular models

In the context of studying bio-nano interactions of silica-based nanomaterials having potential use as adjuvants in immunotherapy and allergens as active pharmaceutical ingredients (APIs) we used the instance map tool to summarise and highlight different workflows generated for investigating immunotoxicity and pharmacologic efficacy endpoints. At the material side, studies focused on silica (SiO_2) nanomaterials in the

size range of 50-100 nm (depending on the method used; TEM, NTA; DLS intensity or number distribution) with different surface modifications, which are reported to be immunologically active in different ways but overall are considered to be safe [46]. In some studies, the impact of material surface (nanotopography) and functional modifications on API binding (molecular initiating event according to the adverse outcome pathway (AOP) concept [47]) were investigated, while in others, the different steps involved in specific immune reaction mechanisms (key events for beneficial vs. adverse outcomes) were analysed. Figure 9 illustrates the comprehensive experimental workflow overarching several immunotoxicity studies highlighting the different routes chosen for these studies, starting with the synthesis, surface functionalisation, and physico-chemical characterisation of the nanomaterials, the bio-nano interaction studies, and the determination of different biological / immunological endpoints.

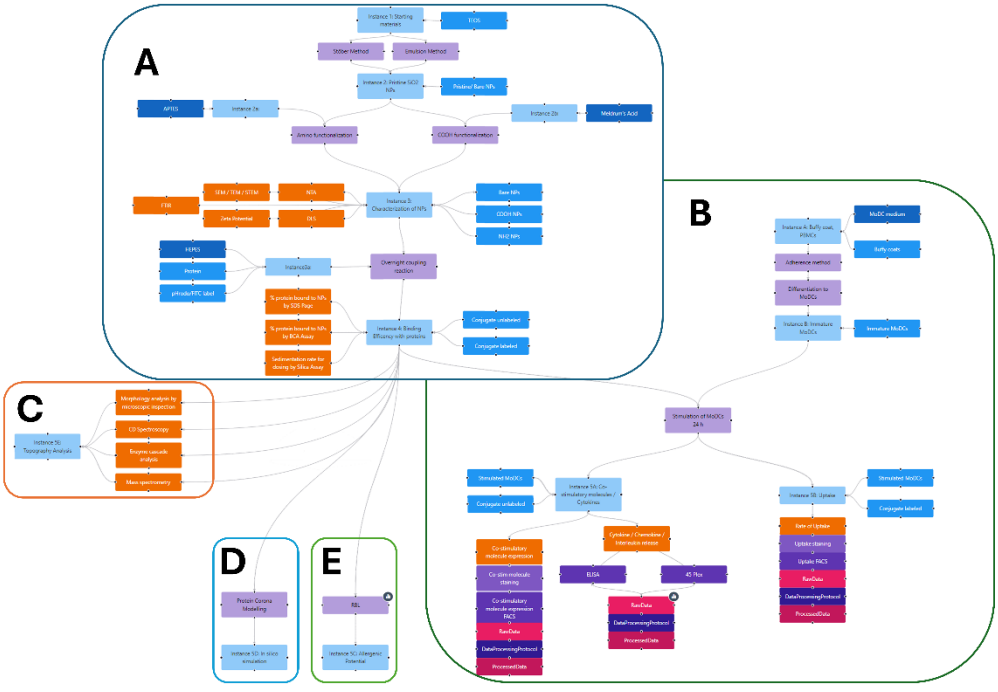

Figure 9: Instance Map showing the immunotoxicity workflow to study the bio-nano interactions of differently functionalized SiO_2 nanomaterials with immune cells. The instance map is divided into Sections A-E, based on the studies of Hasenkopf et al.

[48], Mills-Goodlet et al. [49], Johnson et al. [50], and Punz et. al. [51], highlighting the different approaches and routes that were taken. Section A, which serves as a baseline for all studies, mainly focused on the nanomaterial synthesis and surface modification. The pathway towards more in depth immunological investigations was chosen for Section B, while Sections C & E covered alterations in the protein binding activity depending on the physico-chemical properties provided by chemical surface functionalization and also the observed structural alterations that occurred upon nanomaterial conjugation due to the nanotopography of the materials (mesoporous SiO_2 nanomaterials). Section D depicts the integration of *in silico* predictive modelling approaches with *in vitro* experimental protein corona quantitative and qualitative (epitope rearrangement) determination. The full instance map is accessible at https://instance-maps.stage.sevenpastnine.com for interactive inspection (username: SupportingInfo and password: maps-for-paper).

When studying bio-nano interactions the starting point is typically the synthesis (Or procurement) of the particles, which for SiO_2 nanomaterials is either the Stöber or the emulsion methods followed by chemical modification. Here, nanomaterial functionalisation was realised by addition of amino- or carboxyl- functional groups, with shorter vs. longer aliphatic linkers. Alteration in the particles' nanotopography was realised through pore formation during synthesis using cetyltrimethylammonium bromide. The non-covalent conjugations between nanomaterials and proteins were quantitatively characterised directly by gel electrophoresis and indirectly by quantifying the amount of unbound protein in the supernatant upon several washing steps. *In vitro* Sedimentation, Diffusion and Dosimetry studies were undertaken to determine the cell-delivered dose for all submersed culture conditions based on the specific density and size parameters of the bio-nano conjugates [52]. Finally, comprehensive physico-

chemical characterisation was performed applying a set of analyses according to the reporting standards for bio-nano interactions [43] and (meta)data were uploaded to the NanoCommons Knowledge Base [30] following the principles for data FAIRness and metadata stewardship [29]. These were the necessary baseline requirements to proceed with experiments and were defined as Section A in Figure 9.

The sections concerning the biological and immunological readouts, as well pharmacological efficacy, independently expanded upon Section A. The focus in Figure 9B was on the mechanistic studies following uptake and presentation by professional antigen-presenting cell (APC) models using unmodified SiO_2 nanomaterials [50] vs. differently surface functionalized particles [51]. As a model for APCs, monocyte-derived dendritic cells were generated from human whole blood samples as a preliminary step, again building a BIODA-type of reporting structure following SOPs. Afterwards, these APCs were incubated with the materials generated in Section A and their immunologic activation profile was investigated utilising flow cytometry and ELISA. Figure 9 sections C and E are quite similar in concept [49], and investigated the influence of nanotopography on the protein binding capacity and its impact on epitope integrity, while Johnson et al. 2021 [53] identified that structural alterations of proteins bound to nanomaterials impacted the antigen processing machinery in APCs and could thus impact the outcome in terms of immunomodulation. Here, it should be emphasised that during immunotherapy against type 2 immune diseases, such as in the case of allergy, a shift towards regulatory T cell activation is envisioned. Finally, as depicted in Section D, Hasenkopf et al. 2022 [48] tested the proteins' individual binding efficiencies on differently functionalized SiO_2 nanomaterials using varying conditions, also comparing artificial and real allergen mixtures by applying genuine detection assays suitable for allergenic molecules, *in vitro* and

assessed two recently developed *in silico* protein corona prediction tools against the results from the experimental studies.

The aforementioned studies are complex and comprehensive in themselves and targeted individually to different endpoints. The instance map tool allowed the user to generate large and intertwined workflows referring to multiple research objectives. Whereas a single experiment can already be depicted by an instance map, we herewith displayed their use for visualising integrated batteries of assays, and depicted their applicability as a structural representation of larger collaborative research & development endeavours. Instance maps have thus proven instrumental as a tool for creating and illustrating workflows that combine several sophisticated backgrounds, allowing even less experienced users to capture the bigger picture and still perceive more detailed correlations within a larger context.

Using instance maps for planning and refining data and material workflows in large collaborative projects

As the last examples of application and utilisation of instance maps, their use for reporting of studies with complex workflows and as a study design and materials, samples and data flow planning and tracking tool is presented. The MACRAMÉ project aims to extent the coverage and widen the applicability domains of harmonised OECD test guidelines, OECD guidance documents and international standards (CEN, ISO) by refining existing and developing new advanced physicochemical, human and ecotoxicity characterisation methodologies for market-relevant nanomaterials and the wider group of advanced materials [54] in their complex product matrices. The applicability, relevance and reliability are tested in five industrial use cases. To demonstrate the instance-map-based data management approach of MACRAMÉ, the use case on antibiotics-loaded polymeric nanomaterials is showcased. These

nanomaterials are used for proof-of-concept of treatment for antibiotic-resistant lung bacterial infections. In addition, controls are prepared for imaging purposes to verify the suitability of the MACRAMÉ approach to quantify and characterise the aerosols upon exposure of *in vitro* lung models. Relevant exposure points were identified and used to define the samples that need to be taken from the industrial processes and sent to the characterisation and testing partners. The combinations of advanced materials and complex matrices to be studied include all life-cycle-relevant occurrences of (a) complex product matrices, (b) degraded complex product matrices at the product's end-of-life, (c) regulatory relevant biological matrices for human toxicity testing, (d) environmental matrices for ecotoxicity testing, and relevant forms of the different complex matrices, such as soot and char, and aerosols generated from compounding, machining, use, weathering, degradation or incineration of products.

To achieve such a full characterisation of the materials along their complete life cycle and, at the same time, move the methods forward on their road to standardisation - all in the short runtime of the project - intensive collaboration and unhindered knowledge exchange between all partners is essential. Flows of material and data/information from production to sample preparation (simulating different end-of-life scenarios), to collection of the characterisation data, needs to be organised effectively in order to satisfy the information requirements of downstream experiments and all data needs to be integrated to perform a safety and sustainability evaluation.

The instance map tool was used to visually map out all the exposure points, the characterisation methods applied to these points and the workflows needed to create the materials, the life-cycle samples and to execute the experiments. This ensures that all information required to perform the safety and life-cycle assessment are collected, with all steps documented as part of the planning status (see Figure 10).

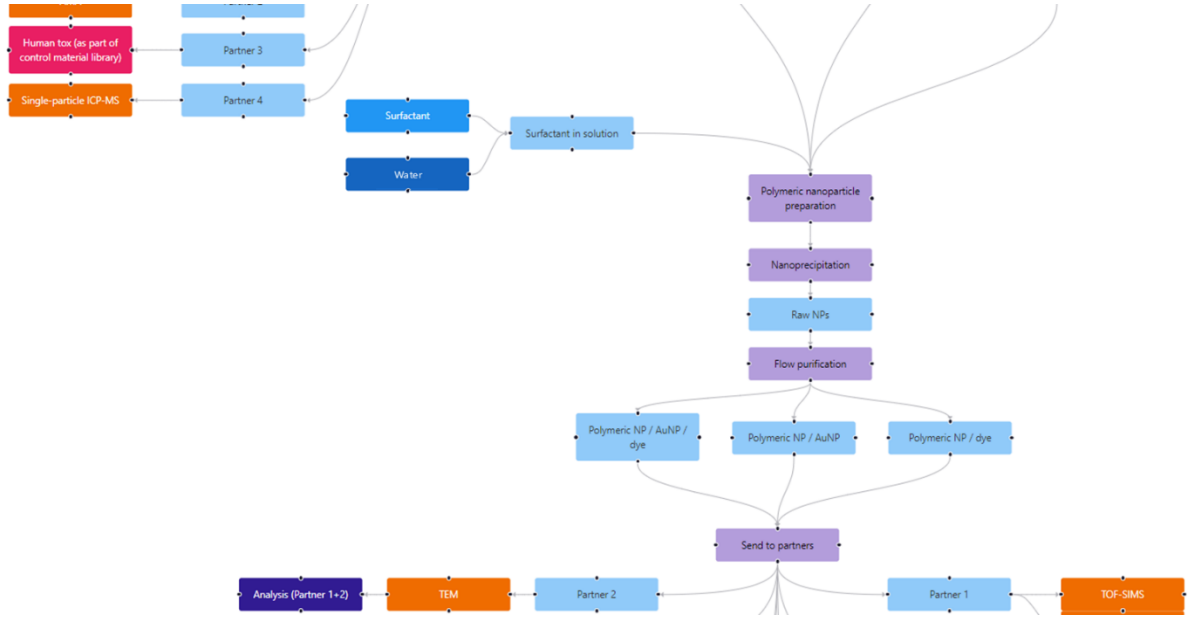

Figure 10: Part of the instance map depicting the planning status of the human and ecotoxicology testing for the MACRAMÉ use case on antibiotics-loaded polymeric nanomaterials. After the production of the loaded nanomaterials, they are sent to many experimental partners performing the different assays. The instance map was crucial to describe the complexity of the workflow, which includes strong cross-partner dependencies like sample preparation by one partner and measurement by another, which must be completed within a specific timeframe. The full instance map is accessible at https://instance-maps.stage.sevenpastnine.com for inspection (username: SupportingInfo and password: maps-for-paper).

Besides nodes representing the materials/samples (instances), characteristics and endpoints to be collected (experimentally and via text/database mining) and nodes describing modification steps applied to the materials and sample (transformation protocols) and testing SOPs (sample preparation, measurement and data processing) as used in the previous examples, the instance maps also focused and clearly defined the chemical and physical treatment and processing steps performed as part of the

manufacturing process, as well as shipping of samples from one partner to another (see Figure 11).

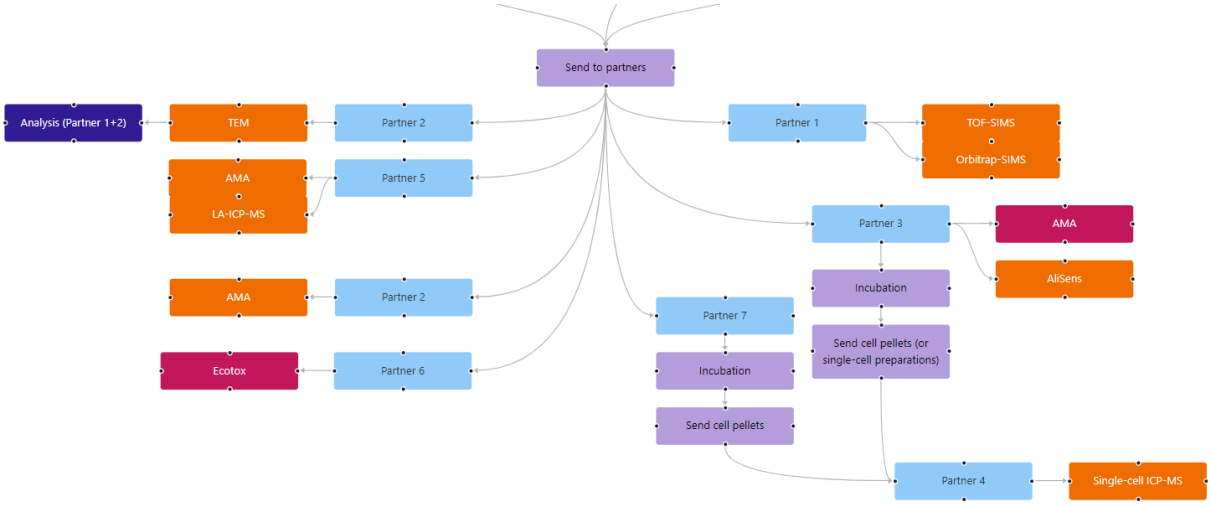

Figure 11: Part of the instance map for the MACRAMÉ use case on antibiotic-loaded nanoparticles representing the shipping of the pristine material to partners performing the human and ecotoxicity testing. The full instance map is accessible at https://instance-maps.stage.sevenpastnine.com for interactive inspection (username: SupportingInfo and password: maps-for-paper).

By adding manufacturing and project management nodes, the instance maps now offer options to collect and document all digitalised information and results produced in upstream tasks of the case studies at one central place for direct (re)use in downstream tasks. In combination with the other components of the MACRAMÉ data management infrastructure and data harmonisation activities, the partners are able to adopt an on-the-fly FAIRification approach [31], in which all research output, including but not limited to sampling plans, study designs, *in vitro* and *in silico* method specifications, protocols, SOPs, and the data created, as well as guidelines, reports, training materials and publications, are directly shared - even in draft versions - by attaching them to instance maps nodes. Since the maps are only available to the

consortium (until / unless they are made public by agreement of all involved parties), they can be used, e.g., to report very detailed partly confidential information on the production processes needed for the life-cycle assessment to evaluate energy and water consumption or as a basis to discuss the amounts of material needed to be shipped to the partners and then the status of the shipment. At the same time, the instance maps are continuously updated to represent more and more the real workflows performed in the use cases with different versions documenting the need and the reasons for deviating from the original planning and when this need became evident. We note here that instance maps could potentially also be utilised to pinpoint where a particular experiment has gone awry or deviated from prior results, including in the case of negative data. This demonstrates that FAIR is not only relevant for secondary reuse of data but that it supports data collection and sharing in large projects from day 1 and that the work invested in FAIRification at the planning stages will immensely reduce the effort for FAIR storage and sharing of data via agreed licences.

Finally, Figure 11 also demonstrates that misusing the colour coding established for clear identification of the node's purpose can be beneficial during planning. Red colours, which are normally used to represent data, were here applied to indicate areas where further discussions were needed on how to perform the experiments or if they are even possible in the time- and budget-frame of the project. Extending the application of instance maps to all the uses described in this paper and potentially many more was only possible by not enforcing strict rules on how different node types can be connected with one other. However, some more guidance might be needed to make the instance maps and the linked data more comparable and interoperable. Now that the applications are better defined and the use cases of the tool have matured, this will be pursued through extending the design guide published in the original instance map paper [28], and by preparing standardised workflows as arrangements

of nodes and/or limiting the way nodes can be connected. To retain some flexibility, the node library could then be extended, for example, by adding specific planning nodes and/or by having customisable nodes.

Key lessons from the implementation cases and future directions

The modifications introduced during the development of the instance map tool, especially the extension of the available node types, opened up many new uses of instance maps for all of the applications presented above, and potentially new ones in the future. However, they also made defining rules on how to connect different nodes less straightforward as in the original approach, where the focus was completely on the fate of the nanomaterial. Newer instance maps are looking into more details of the biological testing system or completely focus on it (for example see Figure 5). Transformation protocol nodes helped to understand which object, e.g., nanomaterial, biological test system or even solvent/medium, underwent a modification but they made the separation between material and medium less obvious. These circumstances also raised the question as to whether materials and medium always have to be associated to an instance or if they can be independent entities when they are used for the first time in a synthesis, functionalisation or exposure protocol (see for example Figures 4 and 11).

Another example, where different groupings of nodes have been used in different applications was the use of properties in combination with sample preparation, measurement and processing protocols and the resulting data. For example, the combination proposed by the ACEnano knowledge base, of sample preparation -> measurement -> raw data -> processing -> processed data, could be placed before or after the node defining the measured property, or could even replacing this node

completely. It was interesting to see how instance maps describing the same study but being created by different users showed significant variations in how nodes were used and connected. This was first recognised when the study from Martinez et al. [32] was used in instance map training and then compared to the original map presented in the publication. Such deviations in instance map design do not cause a problem *per se*. In most cases, it was easy for others to understand the design and flow of the study and to easily identify important results, obviously based on rules of common sense. Only in a few, early cases, the maps needed to be corrected to avoid inconsistencies. The corresponding data, protocols/SOPs and other research outputs could be linked to the maps independently of differences in representation.

To demonstrate that such variations in how instance maps are constructed / nodes are linked can be used to put the focus on different parts of the instance maps, we decided to present all of the examples in the way that the person(s) who performed the experiments had created them, and did not force users to comply with any specific set of rules. However, some more standardisation and a limited set of rules for linking nodes together could speed up comparison (and interoperability) of workflows, one of the main benefits stated during the NIKC curation process. Standardisation would also facilitate the generation of harmonised, comparable data packages combining all information associated with one map, enabling upload of all data to target databases.

There are a number of other areas where the instance maps and the tool could be further extended. The highlighting of specific areas in the maps shown in Figures 3 and 9 were created manually but clearly show that integrating functionalities to create such annotations directly in the tool would be very beneficial. Additionally, better support to link different instance maps or show more detail when hovering over specific parts could reduce the complexity of the maps, especially for complex studies as visualised in Figures 10 and 11 without the need to remove important details. Finally,

the data management and sharing functionality need to be improved to show which information is available and from where, to give access to multiple information sources from one node, and to provide integration with important data management tools like ELNs and protocol repositories. Ways to implement these extensions and improvements are currently under investigation.

It is worth stressing, however, that even if instance maps could drastically change the way data is collected, they are not meant to replace existing data management solutions. Instead, tools implementing the instance map concept should be integrable into existing data ecosystems. Instance maps address two very specific purposes by (1) providing a visual and structured overview of a study, and (2) as an addition to the original concept, linking to resources with additional information for a specific part or component of the study. In this way, they can become the link between different types of personal, institutional, and public information resources (databases and data warehouses, protocol and SOP repositories, software, and source code repositories) and data input and curation services including ELNs. Some ELNs already offer a somewhat similar functionality by allowing user to organise the different steps of the experiments in a workflow. However, as shown in this paper, the instance maps are one level above these workflows since they can represent different levels of detail to show complete, very complex studies, and then zoom into the details of these studies to highlight the metadata and data required at each step. Additionally, different solutions can be used for different types of information customised to the needs of the user and/or community recommendations and not limited to what a specific ELN solution is offering.

Conclusions

From its initial conception as a way to track nanomaterials transformations as reported in literature studies, the instance map approach has undergone very rapid development into a multi-purpose experimental visualisation tool with multiple applications. At its simplest, an instance map can be considered as a graphical abstract summarising the steps in an experimental, computational or combined workflow, demonstrating the materials, their surroundings (medium, environment, organism) and the end-points being measured, and the data flows arising from these at each step of the experiment (synthesis, dispersion, characterisation, exposure, hazard assessment) and/or each stage of the nanomaterials life cycle (production, formulation, application, end of life and disposal or recycling). When applied to standardised regulatory tests or production scenarios, instance maps can be used to provide completeness checks for studies or production batches, ensuring that all necessary parameters to be recorded are captured in the visual model. In this context, instance maps can also be utilised as training tools to emphasise to researchers and operators why specific parameters or checks are essential and to ensure that the complete workflow is understood even where individuals are only responsible for small segments of a workflow. Application of instance mapping at the study design stage can also provide critical insight into bottlenecks and support management aspects such as flows of samples between partners in collaborative research projects, and efforts to support FAIRification of metadata and data prior to data collection save time and resources later.

Creation of the instance mapping software tool described here has greatly enhanced the utility of instance maps, and makes extended applications of instance mapping more accessible, and mapping of highly complex and/or multi-partner

collaborative workflows feasible and practical. The examples presented here highlight the flexibility of the instance mapping software tool, including the capacity for linking of instance maps, and for inclusion of additional category nodes covering quality assurance and quality control, industrial production and management of (planned and actual) materials flows. This flexibility has allowed instance mapping to be used for designing experiments, developing SOPs, and creating and sharing workflows within projects, and as an additional data management tool. However, as the user base expands, the risk of divergent approaches emerging also increases, which will reduce its effectiveness for comparing and integrating datasets. Thus, a balance between flexibility and standardisation will be implemented, through development guiding principles for the design of instance maps and the optimal connection of nodes, to maximise its potential for harmonization and standardisation purposes. This would facilitate the generation of harmonised, comparable data packages combining all information associated with one map, enabling upload of all data to target databases such as the NanoCommons Knowledge Base. Integration of the instance map tool with other data management solutions, such as electronic laboratory notebooks, protocols registries and databases will further enhance its utility and position it as a key FAIR Enabling Resource for safety and sustainability assessment of nanoscale and advanced materials, and beyond.

Supporting Information

All instance maps created in the new instance map tool are available from https://instance-maps.stage.sevenpastnine.com (username: SupportingInfo and password: maps-for-paper).

Funding

This work was funded by the EU Horizon 2020 projects NanoCommons (Grant Agreement No. 731032), NanoFASE (Grant Agreement No. 646002) and CompSafeNano (Grant Agreement No. 101008099), and the Horizon Europe projects WorldFAIR (Grant Agreement No. 101058393), MACRAMÉ (Grant Agreement No. 101092686), and PINK (Grant Agreement No. 101137809). Additional support came from the SmartCERIALS project of the Austrian Research Promotion Agency (FFG, Grant No. 890610), the Innovate UK for UoB participation in WorldFAIR (Grant No. 1831977) and MACRAMÉ (Grant No. 10066165). This material is also based upon work supported by the NSF and the EPA through the Center for the Environmental Implications of NanoTechnology (CEINT), and the INFRAMES network funded through NSF's AccelNet program Award 2114682. Graphical abstract was created using BioRender.com (accessed on 20 March 2024).

References

(1) Sudha, P. N.; Sangeetha, K.; Vijayalakshmi, K.; Barhoum, A. Chapter 12 - Nanomaterials History, Classification, Unique Properties, Production and Market. In *Emerging Applications of Nanoparticles and Architecture Nanostructures*; Barhoum, A., Makhlouf, A. S. H., Eds.; Micro and Nano Technologies; Elsevier, 2018; pp 341–384. doi:10.1016/B978-0-323-51254-1.00012-9

(2) Furxhi, I.; Costa, A.; Vázquez-Campos, S.; Fito-López, C.; Hristozov, D.; Ramos, J. A. T.; Resch, S.; Cioffi, M.; Friedrichs, S.; Rocca, C.; Valsami-Jones, E.; Lynch, I.; Araceli, S. J.; Farcal, L. *RSC Sustain.* **2023**, *1*, 234–250. doi:10.1039/D2SU00101B

(3) Ashik, U. P. M.; Viswan, A.; Kudo, S.; Hayashi, J. Chapter 3 - Nanomaterials as Catalysts. In *Applications of Nanomaterials*; Mohan Bhagyaraj, S., Oluwafemi, O. S., Kalarikkal, N., Thomas, S., Eds.; Micro and Nano Technologies; Woodhead Publishing, 2018; pp 45–82. doi:10.1016/B978-0-08-101971-9.00003-X

(4) Saeed, A.; Munir, S.; Gull, N.; Khan, S. M. 15 - Nanomaterials for Carbon Capture and Their Conversion to Useful Products for Sustainable Energy Production. In *Nanomaterials in Biomass Conversion*; Rizwan, K., Bilal, M., Eds.;

Woodhead Series in Bioenergy; Woodhead Publishing, 2024; pp 369–395. doi:10.1016/B978-0-443-13500-2.00015-8
(5) Shellaiah, M.; Sun, K. W. *Chemosensors* **2020**, *8*, 55. doi:10.3390/chemosensors8030055
(6) Sui, X.; Downing, J. R.; Hersam, M. C.; Chen, J. *Materials Today* **2021**, *48*, 135–154. doi:10.1016/j.mattod.2021.02.001
(7) Lin, W. *Chem. Rev.* **2015**, *115*, 10407–10409. doi:10.1021/acs.chemrev.5b00534
(8) Bardhan, N. *MRS Communications* **2022**, *12*, 1119–1139. doi:10.1557/s43579-022-00257-7
(9) Lowry, G. V.; Avellan, A.; Gilbertson, L. M. *Nat. Nanotechnol.* **2019**, *14*, 517–522. doi:10.1038/s41565-019-0461-7
(10) Zhang, P.; Lynch, I.; Handy, R. D.; White, J. C. 1 - A Brief History of Nanotechnology in Agriculture and Current Status. In *Nano-Enabled Sustainable and Precision Agriculture*; Zhang, P., Lynch, I., White, J. C., Handy, R. D., Eds.; Academic Press, 2023; pp 3–14. doi:10.1016/B978-0-323-91233-4.00002-8
(11) Walczyk, D.; Bombelli, F. B.; Monopoli, M. P.; Lynch, I.; Dawson, K. A. *J. Am. Chem. Soc.* **2010**, *132*, 5761–5768. doi:10.1021/ja910675v
(12) Wheeler, K. E.; Chetwynd, A. J.; Fahy, K. M.; Hong, B. S.; Tochihuitl, J. A.; Foster, L. A.; Lynch, I. *Nat. Nanotechnol.* **2021**, *16*, 617–629. doi:10.1038/s41565-021-00924-1
(13) Lowry, G. V.; Gregory, K. B.; Apte, S. C.; Lead, J. R. *Environ. Sci. Technol.* **2012**, *46*, 6893–6899. doi:10.1021/es300839e
(14) Svendsen, C.; Walker, L. A.; Matzke, M.; Lahive, E.; Harrison, S.; Crossley, A.; Park, B.; Lofts, S.; Lynch, I.; Vázquez-Campos, S.; Kaegi, R.; Gogos, A.; Asbach, C.; Cornelis, G.; von der Kammer, F.; van den Brink, N. W.; Mays, C.; Spurgeon, D. J. *Nat. Nanotechnol.* **2020**, *15*, 731–742. doi:10.1038/s41565-020-0742-1
(15) Rauscher, H.; Rasmussen, K.; Sokull-Klüttgen, B. *Chemie Ingenieur Technik* **2017**, *89*, 224–231. doi:10.1002/cite.201600076
(16) Clausen, L. P. W.; Hansen, S. F. *Nature Nanotech* **2018**, *13*, 766–768. doi:10.1038/s41565-018-0256-2
(17) Lowry, G. V.; Hill, R. J.; Harper, S.; Rawle, A. F.; Hendren, C. O.; Klaessig, F.; Nobbmann, U.; Sayre, P.; Rumble, J. *Environ. Sci.: Nano* **2016**, *3*, 953–965. doi:10.1039/C6EN00136J
(18) Lynch, I.; Weiss, C.; Valsami-Jones, E. *Nano Today* **2014**, *9*, 266–270. doi:10.1016/j.nantod.2014.05.001
(19) Izak-Nau, E.; Huk, A.; Reidy, B.; Uggerud, H.; Vadset, M.; Eiden, S.; Voetz, M.; Himly, M.; Duschl, A.; Dusinska, M.; Lynch, I. *RSC Adv.* **2015**, *5*, 84172–84185. doi:10.1039/C5RA10187E
(20) Hendren, C.O., Lowry, G.V., Unrine, J.M. and Wiesner, M.R., 2015. A functional assay-based strategy for nanomaterial risk forecasting. SOTEN, 536, pp.1029-1037. doi:10.1016/j.scitotenv.2015.06.100
(21) Baer, D. R.; Munusamy, P.; Thrall, B. D. *Biointerphases* **2016**, *11*, 04B401. doi:10.1116/1.4964867
(22) Mülhopt, S.; Diabaté, S.; Dilger, M.; Adelhelm, C.; Anderlohr, C.; Bergfeldt, T.; Gómez de la Torre, J.; Jiang, Y.; Valsami-Jones, E.; Langevin, D.; Lynch, I.; Mahon, E.; Nelissen, I.; Piella, J.; Puntes, V.; Ray, S.; Schneider, R.; Wilkins, T.; Weiss, C.; Paur, H.-R. *Nanomaterials* **2018**, *8*, 311. doi:10.3390/nano8050311
(23) Rumble, J.; Freiman, S.; Teague, C. *Chemistry International* **2015**, *37*, 3–7. doi:10.1515/ci-2015-0402

(24) van Rijn, J.; Afantitis, A.; Culha, M.; Dusinska, M.; Exner, T. E.; Jeliazkova, N.; Longhin, E. M.; Lynch, I.; Melagraki, G.; Nymark, P.; Papadiamantis, A. G.; Winkler, D. A.; Yilmaz, H.; Willighagen, E. *Journal of Cheminformatics* **2022**, *14*, 57. doi:10.1186/s13321-022-00614-7
(25) Chetwynd, A. J.; Wheeler, K. E.; Lynch, I. *Nano Today* **2019**, *28*, 100758. doi:10.1016/j.nantod.2019.06.004
(26) Rasmussen, K.; Rauscher, H.; Kearns, P.; González, M.; Riego Sintes, J. *Regulatory Toxicology and Pharmacology* **2019**, *104*, 74–83. doi:10.1016/j.yrtph.2019.02.008
(27) Amos, J. D.; Tian, Y.; Zhang, Z.; Lowry, G. V.; Wiesner, M. R.; Hendren, C. O. *NanoImpact* **2021**, *23*, 100331. doi:10.1016/j.impact.2021.100331
(28) Amos, J. D.; Zhang, Z.; Tian, Y.; Lowry, G. V.; Wiesner, M. R.; Hendren, C. O. *Sci Data* **2024**, *11*, 173. doi:10.1038/s41597-024-03006-8
(29) Papadiamantis, A. G.; Klaessig, F. C.; Exner, T. E.; Hofer, S.; Hofstaetter, N.; Himly, M.; Williams, M. A.; Doganis, P.; Hoover, M. D.; Afantitis, A.; Melagraki, G.; Nolan, T. S.; Rumble, J.; Maier, D.; Lynch, I. *Nanomaterials* **2020**, *10*, 2033. doi:10.3390/nano10102033
(30) Maier, D.; Exner, T. E.; Papadiamantis, A. G.; Ammar, A.; Tsoumanis, A.; Doganis, P.; Rouse, I.; Slater, L. T.; Gkoutos, G. V.; Jeliazkova, N.; Ilgenfritz, H.; Ziegler, M.; Gerhard, B.; Kopetsky, S.; Joshi, D.; Walker, L.; Svendsen, C.; Sarimveis, H.; Lobaskin, V.; Himly, M.; van Rijn, J.; Winckers, L.; Millán Acosta, J.; Willighagen, E.; Melagraki, G.; Afantitis, A.; Lynch, I. *Front. Phys.* **2023**, *11*. doi:10.3389/fphy.2023.1271842
(31) Exner, T. E.; Papadiamantis, A. G.; Melagraki, G.; Amos, J. D.; Bossa, N.; Gakis, G. P.; Charitidis, C. A.; Cornelis, G.; Costa, A. L.; Doganis, P.; Farcal, L.; Friedrichs, S.; Furxhi, I.; Klaessig, F. C.; Lobaskin, V.; Maier, D.; Rumble, J.; Sarimveis, H.; Suarez-Merino, B.; Vázquez, S.; Wiesner, M. R.; Afantitis, A.; Lynch, I. *Frontiers in Physics* **2023**, *11*
(32) Martinez, D. S. T.; Da Silva, G. H.; de Medeiros, A. M. Z.; Khan, L. U.; Papadiamantis, A. G.; Lynch, I. *Nanomaterials* **2020**, *10*, 1936. doi:10.3390/nano10101936
(33) Maia, M. T.; Delite, F. S.; da Silva, G. H.; Ellis, L.-J. A.; Papadiamantis, A. G.; Paula, A. J.; Lynch, I.; Martinez, D. S. T. *Journal of Hazardous Materials* **2024**, *461*, 132623. doi:10.1016/j.jhazmat.2023.132623
(34) Levard, C.; Reinsch, B. C.; Michel, F. M.; Oumahi, C.; Lowry, G. V.; Brown, G. E. Jr. *Environ. Sci. Technol.* **2011**, *45*, 5260–5266. doi:10.1021/es2007758
(35) Levard, C.; Hotze, E. M.; Lowry, G. V.; Brown, G. E. Jr. *Environ. Sci. Technol.* **2012**, *46*, 6900–6914. doi:10.1021/es2037405
(36) Lynch, I.; Afantitis, A.; Exner, T.; Himly, M.; Lobaskin, V.; Doganis, P.; Maier, D.; Sanabria, N.; Papadiamantis, A. G.; Rybinska-Fryca, A.; Gromelski, M.; Puzyn, T.; Willighagen, E.; Johnston, B. D.; Gulumian, M.; Matzke, M.; Green Etxabe, A.; Bossa, N.; Serra, A.; Liampa, I.; Harper, S.; Tämm, K.; Jensen, A. C.; Kohonen, P.; Slater, L.; Tsoumanis, A.; Greco, D.; Winkler, D. A.; Sarimveis, H.; Melagraki, G. *Nanomaterials* **2020**, *10*, 2493. doi:10.3390/nano10122493
(37) Silva, P. V.; Silva, A. R. R.; Clark, N. J.; Vassallo, J.; Baccaro, M.; Medvešček, N.; Grgić, M.; Ferreira, A.; Busquets-Fité, M.; Jurkschat, K.; Papadiamantis, A. G.; Puntes, V.; Lynch, I.; Svendsen, C.; van den Brink, N. W.; Handy, R. D.; van Gestel, C. A. M.; Loureiro, S. *Science of The Total Environment* **2023**, *873*, 162160. doi:10.1016/j.scitotenv.2023.162160
(38) Khodaparast, Z.; van Gestel, C. A. M.; Silva, A. R. R.; Cornelis, G.; Lahive, E.; Etxabe, A. G.; Svendsen, C.; Baccaro, M.; van den Brink, N.; Medvešček, N.;

Novak, S.; Kokalj, A. J.; Drobne, D.; Jurkschat, K.; Loureiro, S. *NanoImpact* **2023**, *29*, 100454. doi:10.1016/j.impact.2023.100454
(39) Lahive, E.; Schultz, C. l.; Van Gestel, C. a. m.; Robinson, A.; Horton, A. a.; Spurgeon, D. j.; Svendsen, C.; Busquets-Fité, M.; Matzke, M.; Green Etxabe, A. *Environmental Toxicology and Chemistry* **2021**, *40*, 1859–1870. doi:10.1002/etc.5031
(40) Clark, N.; Vassallo, J.; Silva, P. V.; Silva, A. R. R.; Baccaro, M.; Medvešček, N.; Grgić, M.; Ferreira, A.; Busquets-Fité, M.; Jurkschat, K.; Papadiamantis, A. G.; Puntes, V.; Lynch, I.; Svendsen, C.; van den Brink, N. W.; van Gestel, C. A. M.; Loureiro, S.; Handy, R. D. *Science of The Total Environment* **2022**, *850*, 157912. doi:10.1016/j.scitotenv.2022.157912
(41) OECD. *Test No. 202: Daphnia Sp. Acute Immobilisation Test*; Organisation for Economic Co-operation and Development: Paris, 2004
(42) Sebastiani, M.; Charitidis, C.; Koumoulos, E. P. Main Introduction to the CHADA Concept and Case Studies, 2019. doi:10.5281/zenodo.2636609
(43) Faria, M.; Björnmalm, M.; Thurecht, K. J.; Kent, S. J.; Parton, R. G.; Kavallaris, M.; Johnston, A. P. R.; Gooding, J. J.; Corrie, S. R.; Boyd, B. J.; Thordarson, P.; Whittaker, A. K.; Stevens, M. M.; Prestidge, C. A.; Porter, C. J. H.; Parak, W. J.; Davis, T. P.; Crampin, E. J.; Caruso, F. *Nature Nanotech* **2018**, *13*, 777–785. doi:10.1038/s41565-018-0246-4
(44) Adamovic, N.; Boskovic, B.; Celuch, M.; Charitidis, C.; Friis, J.; Goldbeck, G.; Hashibon, A.; Hurtós, E.; Sebastiani, M.; Simperler, A. *Report on Advanced Materials Modelling and Characterisation: Strategies for Integration and Interoperability*; Zenodo, 2021. doi:10.5281/zenodo.4912683
(45) Krebs, A.; van Vugt-Lussenburg, B. M. A.; Waldmann, T.; Albrecht, W.; Boei, J.; ter Braak, B.; Brajnik, M.; Braunbeck, T.; Brecklinghaus, T.; Busquet, F.; Dinnyes, A.; Dokler, J.; Dolde, X.; Exner, T. E.; Fisher, C.; Fluri, D.; Forsby, A.; Hengstler, J. G.; Holzer, A.-K.; Janstova, Z.; Jennings, P.; Kisitu, J.; Kobolak, J.; Kumar, M.; Limonciel, A.; Lundqvist, J.; Mihalik, B.; Moritz, W.; Pallocca, G.; Ulloa, A. P. C.; Pastor, M.; Rovida, C.; Sarkans, U.; Schimming, J. P.; Schmidt, B. Z.; Stöber, R.; Strassfeld, T.; van de Water, B.; Wilmes, A.; van der Burg, B.; Verfaillie, C. M.; von Hellfeld, R.; Vrieling, H.; Vrijenhoek, N. G.; Leist, M. *Arch Toxicol* **2020**, *94*, 2435–2461. doi:10.1007/s00204-020-02802-6
(46) Navarro-Tovar, G., Palestino G., Rosales-Mendoza, S. An overview on the role of silica-based materials in vaccine development. Expert Review of Vaccines 15(11): 1449-1462 (2016). doi: 10.1080/14760584.2016.1188009
(47) Hofer, S.; Hofstätter, N.; Punz, B.; Hasenkopf, I.; Johnson, L.; Himly, M. *WIREs Nanomedicine and Nanobiotechnology* **2022**, *14*, e1804. doi:10.1002/wnan.1804
(48) Hasenkopf, I.; Mills-Goodlet, R.; Johnson, L.; Rouse, I.; Geppert, M.; Duschl, A.; Maier, D.; Lobaskin, V.; Lynch, I.; Himly, M. *Nano Today* **2022**, *46*, 101561. doi:10.1016/j.nantod.2022.101561
(49) Mills-Goodlet, R.; Johnson, L.; Hoppe, I. J.; Regl, C.; Geppert, M.; Schenck, M.; Huber, S.; Hauser, M.; Ferreira, F.; Hüsing, N.; Huber, C. G.; Brandstetter, H.; Duschl, A.; Himly, M. *Nanoscale* **2021**, *13*, 20508–20520. doi:10.1039/D1NR05958K
(50) Johnson, L.; Aglas, L.; Punz, B.; Dang, H.-H.; Christ, C.; Pointner, L.; Wenger, M.; Hofstaetter, N.; Hofer, S.; Geppert, M.; Andosch, A.; Ferreira, F.; Horejs-Hoeck, J.; Duschl, A.; Himly, M. *Nanoscale* **2023**, *15*, 2262–2275. doi:10.1039/D2NR05181H

(51) Punz, B.; Johnson, L.; Geppert, M.; Dang, H.-H.; Horejs-Hoeck, J.; Duschl, A.; Himly, M. *Pharmaceutics* **2022**, *14*, 1103. doi:10.3390/pharmaceutics14051103
(52) Himly, M.; Geppert, M.; Hofer, S.; Hofstätter, N.; Horejs-Höck, J.; Duschl, A. *Small* **2020**, *16*, 1907483. doi:10.1002/smll.201907483
(53) Johnson, L.; Aglas, L.; Soh, W. T.; Geppert, M.; Hofer, S.; Hofstätter, N.; Briza, P.; Ferreira, F.; Weiss, R.; Brandstetter, H.; Duschl, A.; Himly, M. *International Journal of Molecular Sciences* **2021**, *22*, 10895. doi:10.3390/ijms221910895
(54) Advanced Materials for Industrial Leadership - European Commission https://research-and-innovation.ec.europa.eu/research-area/industrial-research-and-innovation/key-enabling-technologies/chemicals-and-advanced-materials/advanced-materials-industrial-leadership_en (accessed Mar 19, 2024)