Communication

European Registry of Materials: global, unique identifiers for (undisclosed) nanomaterials

Jeaphianne van Rijn[1, orcid:0000-0001-5026-7705,*], **Antreas Afantitis**[2, orcid: 0000-0002-0977-8180], **Mustafa Culha**[3, orcid:0000-0002-3844-5190], **Maria Dusinska**[4, orcid:0000-0003-1358-1652], **Thomas Exner**[5, orcid:0000-0002-1849-5246], **Nina Jealiazkova**[6, orcid:0000-0002-4322-6179], **Eleonora Marta Longhin**[4, orcid:0000-0002-3939-7135], **Iseult Lynch**[7, orcid: 0000-0003-4250-4584], **Georgia Melagraki**[2, orcid: 0000-0001-7547-2342], **Penny Nymark**[8, orcid:0000-0002-3435-7775], **Anastasios G. Papadiamantis**[2,7, orcid: 0000-0002-1297-3104], **David A. Winkler**[9,10,11,, orcid: 0000-0002-7301-6076], **Hulya Yilmaz**[3, orcid: 0000-0003-4592-6432], **and Egon Willighagen**[1,orcid:0000-0001-7542-0286]

1 Dept of Bioinformatics – BiGCaT, NUTRIM, FHML, Maastricht University, The Netherlands
2 NovaMechanics Ltd., 1070, Nicosia, Cyprus
3 Sabanci University Nanotechnology Research and Application Center (SUNUM), Tuzla, 34956 Istanbul, Turkey
4 Health Effects Laboratory, Department of Environmental Chemistry NILU, Norwegian Institute for Air Research, 2007 Kjeller, Norway
5 Seven Past Nine d.o.o., 1380 Cerknica, Slovenia
6 Ideaconsult Ltd., Sofia 1000, Bulgaria
7 School of Geography, Earth and Environmental Sciences, University of Birmingham, Edgbaston, B15 2TT, United Kingdom
8 Institute of Environmental Medicine, Karolinska Institute, Stockholm, Sweden
9 School of Biochemistry and Genetics, La Trobe Institute for Molecular Science, La Trobe University, Bundoora, Australia
10 Monash University, Parkville, Australia
11 School of Pharmacy, University of Nottingham, Nottingham, United Kingdom

* Correspondence: j.vanrijn@maastrichtuniversity.nl

Abstract: Data management of nanomaterials and nanosafety data needs to operate under the FAIR (findability, accessibility, interoperability, and reusability) principles and this requires a unique, global identifier per nanomaterial. Existing identifiers may not always be applicable, or sufficient to definitively identify the specific nanomaterial utilised in a particular study, resulting in the use of textual descriptions in research project communications and reporting. To ensure that internal project documentation can later be linked to publicly released data and knowledge for the specific nanomaterials, or even to specific batches and variants of nanomaterials utilised in that project, a new identifier is proposed: the European Registry of Materials Identifier. We here describe the background to this new identifier, including FAIR interoperability as defined by FAIRSharing, identifiers.org and the CHEMINF ontology, how it complements other identifiers such as CAS numbers and the ongoing efforts to extend the InChI identifier to cover nanomaterials, and provide examples of its use in various H2020-funded nanosafety projects.

Keywords: nanomaterial, FAIR, identifier

---

1. Introduction

Management of nanomaterial and nanosafety data is essential for development of a transparent risk governance model in which decisions are based on high quality datasets that must be appropriately cited [1,2]. Clearly, this management of nanosafety data requires a significant initial effort, but, in the medium term, the enhanced re-usability of nanosafety data will save the European community a significant amount of money [3]. Recently, the FAIR approach for data findability, accessibility, interoperability, and reusability was formulated to optimize the utility of decades of research on these aspects [4,5]. A central element of the FAIR resource output is the global, unique identifier, the workhorse of data integration for many years.

Examples of such identifiers for chemicals include the Chemical Abstracts Service (CAS) registry number [6] and the International Union of Pure and Applied Chemistry (IUPAC) International Chemical Identifier or InChI [7]. The first is assigned to all chemical substances after publication and indexing, while the second can be computed according to an agreed set of rules and abstractions to represent chemical structures. However, these two identifiers are still not sufficient for the wide range of nanosafety research projects being undertaken. The InChI is based on chemical graphs and cannot fully describe the chemistry of engineered nanomaterials (ENMs), although a draft extension of InChI for nanomaterials (NInChI) has been defined recently [8]. The goal of the NInChI is to describe core features of nanomaterials (from their centre towards their surface, the aspect most closely linked to their interactions and potential toxicity). Its purpose is to support integration of datasets of similar materials, and as such is considered a group identifier. CAS registry numbers cannot be created by researchers themselves. This is an issue for internal communication and reporting in (European) research projects where the full physicochemical properties are often measured, but not always reported outside the project, with scientific publications often coming well after the project has ended. Thus, to meet the FAIR requirements, we need a complementary global, unique identifier for unpublished or internally published communication and reporting that can be used without disclosing sensitive or embargoed information but which allows linking of knowledge in a clearly defined way, supporting our nanosafety research and development [9].

Here we describe a recent initiative of several EU NanoSafety Cluster projects to establish such an identifier for use within projects and beyond, modelled on similar efforts in the pharmaceutical industry. In the case of pharmaceuticals, chemicals reported by a company receive an identifier, without the need to disclose their structures [10]. This identifier can be used in, for example, clinical trials and publications, thus linking data to the unique registry number for the specific chemical as information becomes public [11]. Applying this approach to nanomaterials and nanosafety research, the projects compile the list of nanomaterials that they will use and “register” these on behalf of the project. Each nanomaterial then receives an identifier, that may be accompanied by additional information such as supplier, composition, size and batch number, or can be simply a placeholder that does not disclose their structure or composition. This identifier can be used in, for example, project (deliverable) reports and subsequent publications. At some point the chemical structure is disclosed and the identifier is linked to the identity of the material. While this expectation is not explicitly required, releasing data is the default for European Horizon 2020 projects.

We here report the development of a public registry where nanomaterial identifiers can be registered, the European Registry of Materials (nanocommons.github.io/identifiers/). The project is currently governed by the European Horizon 2020 NanoCommons project and the registry itself is available under a CCZero license, removing obstacles for reuse.

2. Materials and Methods

The registry contains supporting pages written in Markdown [12] hosted on GitHub (github.com/nanocommons/identifiers) and a central registry of materials in Turtle [13] (github.com/NanoCommons/identifiers/blob/master/registry). The webpage was autogenerated by GitHub Pages. The European Registry of Materials Identifier was registered with FAIRSharing [14] and identifiers.org [15]. The latter provided us with an official compact identifier structure [16]. The registry Turtle uses the RDF Schema specification for storing labels and types of nanomaterials as *chemical substances*, using the `CHEBI_59999` term from the ChEBI ontology [17].

3. Results

3.1. The European Registry of Materials Identifier

We established the European Registry of Materials (nanocommons.github.io/identifiers/) in April 2019, allowing us to create global, unique identifiers to track nanomaterials used in individual research projects with minimal required information upfront. This provides users with full control over when to disclose or otherwise provide additional information on specific nanomaterials and accompanying physico-chemical and toxicological data. It allows users to discuss nanomaterials in a FAIR way, even before the detailed chemistry of the nanomaterials has been established. At the time of writing, six nanosafety projects have registered materials with the registry: NanoSolveIT, NanoFASE, RiskGONE, NanoTest, caLIBRAte, and SbD4Nano. Additionally, the European

Gov4Nano project implemented these identifiers in several data reuse case studies, though no materials have been registered at the time of writing this paper.

The total number of European Registry of Materials Identifiers (or ERM identifiers for short) for registered materials is currently over 300. The European Registry of Materials dentifier is registered with FAIRSharing at fairsharing.org/bsg-s001384/. The registration of the identifier in identifiers.org is available at registry.identifiers.org/registry/erm (MIR:00000763). This also defined the 'erm' namespace for the compact identifier. Finally, the ERM identifier has also been added to the Chemical Information Ontology (CHEMINF) [18] and is included in the eNanoMapper ontology [19], allowing encoding of the ERM identifiers in Resource Description Framework representations of nanosafety data and annotation of ERM identifiers in spreadsheets and ISA-Tab files [20].

As an example, the NanoSolveIT project registered one material with the `ERM00000001` identifier. The full Uniform Resource Identifier (URI) for this compound is `https://nanocommons.github.io/identifiers/registry#ERM00000001` which is too long to be used in documentation. The corresponding compact identifier `erm:ERM00000001` is easy to use in written material, analogous to the use of Protein Data Bank (PDB) identifiers for proteins in journals.

3.2. Registering identifiers

Registration of new identifiers ultimately occurs by adding the new identifier and a label from the project assigning the identifier. The label can be a material name, but also something generic such as "nanomaterial 1". It is the minimally required information about a registered material. Eventually a lot more information should become available about each registered identifier. Currently, the registry entry for one of the identifiers in Turtle format is: .

```
erm:ERM00000001 a obo:CHEBI_59999; rdfs:label "NanoSolveIT Material 1".
```

Other information that can be provided includes a parent ERM identifier, the chemical composition, a batch number, an ontology material classification (using the eNanoMapper ontology), a web page, and a provider, contact, or project name. If an ontology classification does not currently exist for a specific nanomaterial, users are able to request it through the eNanoMapper ontology GitHub page github.com/enanomapper/ontologies/issues/new/choose [19].

Practically, ERM identifiers for materials can be registered in two ways:

i) by creating a request in the GitHub project issue tracker, as explained in a dedicated tutorial at nanocommons.github.io/identifiers/register; or

ii) by communication (e.g. email or pull requests) to the NanoCommons team, which is the way most current identifiers have been registered. The second approach works best when multiple identifiers need to be registered together, e.g. at the initial stages of a new project when decisions around the nanomaterials to be utilised in the project are finalised.

3.3. Use of ERM identifiers by EU NanoSafety Cluster projects

The EU Nanosafety Cluster (NSC, www.nanosafetycluster.eu/) is a self-organising group of projects addressing various aspects of nanomaterials safety that has been operating since the start of the 7th Framework Programme [21]. Its exact composition varies as projects finish and new ones begin, but its collaborative activities are maintained through a set of overarching working groups, including WG-F on data management which supports the projects in aspects of FAIR data including promoting the ERM. This section describes some of the current NSC projects (in no particular order) that have registered ERM identifiers, how they are using them, and projects that are actively involved with supporting community adoption of the identifier.

3.3.1 NanoCommons

NanoCommons (nanocommons.eu/) is the nanosafety infrastructure project whose goal is to build a common data management environment across European and international nanosafety projects. It is not directly producing data itself and, thus, ERM identifiers have not been registered. However, within its role as a cross-project enabler it accepted the responsibility to develop the ERM and host it on the NanoCommons Github. Additional community support activities undertaken by NanoCommons include extending the common ontology, the eNanoMapper ontology, which has been maintained by NanoCommons in recent years [22]. The ERM identifier aligns well with this goal as it allows nanosafety research projects to globally and uniquely refer to specific nanomaterials. NanoCommons is also driving the development of the NInChI which aims to provide the means by which specific information related to individual nanomaterials in the ERM can be added, and similarity of materials in the registry can be assessed.

Furthermore, via its Transnational Access activities (www.nanocommons.eu/ta-access/) NanoCommons was responsible for guiding the data management processes of the NanoFASE project. In this way, we were able to promote the use of ERMs for the nanomaterials used within the project. The use of ERMs as a unique identifier for the project's nanomaterials was included, as a provision, in the revision and updating of the adopted Nanoinformatics Knowledge Commons (NIKC, ceint.duke.edu/research/nikc, [23]) data curation templates (see Figure 1) and relevant guidance was included in the guides developed through the collaboration of the NanoFASE and NanoCommons projects [24], including the international partners Center for the Environmental Implications of Nanomaterials (CEINT) of Duke University and Team Helium LLC (U.S.A.).

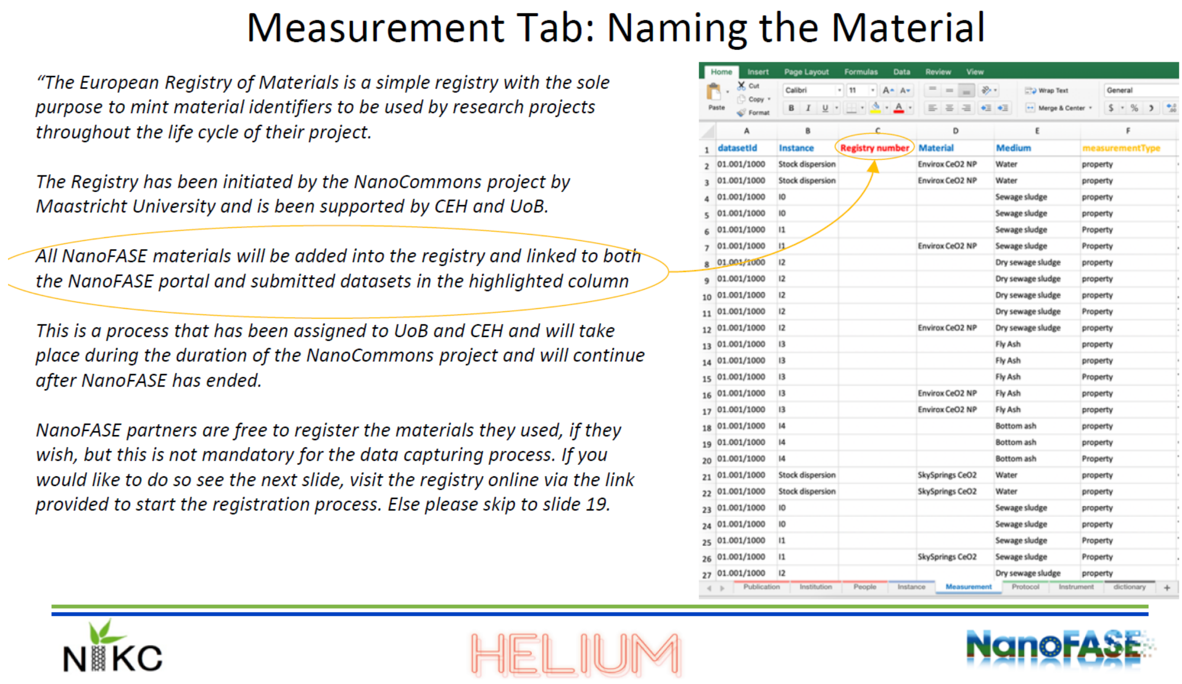

Figure 1. Implementation of the ERM identifiers in the revised NIKC templates used in the NanoFASE project and inclusion in the guidelines document.

3.3.2. RiskGONE

RiskGONE (riskgone.eu/) has registered twelve ERM identifiers for the nanomaterials selected for use in the round robin validation exercises for human and environmental health assay pre-validation. Within the project, these identifiers are linked to a label (currently "RiskGONE NM 1-12"), supplier, supplier code, and batch or vial number. This information is currently kept in a private repository for internal use, and stored as a versioned Markdown file, with Word and PDF exports for convenience. The ERM identifiers are used in the project`s documents, deliverables and reports and are integrated in the eNanoMapper database data entry workflow [1,2]. The RiskGONE data entry templates are following the current Nanosafety Cluster practice of using Excel templates for data logging. Physico-chemical characterization data generated by project partners is entered into a NANoREG style template for physchem properties [25,26] and data from *in vitro* toxicity tests for human and eco-toxicological hazard assessment is entered into adapted IOM–Nano–EHS data templates [1], which have been developed over the last decade and used by a number of EU FP7 and H2020 projects. All templates are distributed through an online Template Wizard tool (search.data.enanomapper.net/projects/riskgone/datatemplates/), and include a dedicated "Materials" sheet, which is automatically filled in with the ERM identifiers, names, types, and other details of the materials used in the project. This spreadsheet serves as a lookup for selecting the materials on data entry and ensures consistent use of material identifiers across the project.

An additional 196 ERM identifiers are used in a data extraction project in collaboration with the SbD4Nano project (www.sbd4nano.eu/). Data was extracted for all nanomaterials mentioned in 25 papers that were previously analyzed for adverse outcomes [27]. This resulted in 196 nanomaterials with matching ERM identifiers. The nanomaterials and their Molecular Initiating Events (MIEs) are captured in an RDF schema and annotated using standard metadata vocabularies and specialized ontologies (manuscript in preparation). This way, the identifiers are linked to a nanomaterial name, the reference from which the data is extracted, the PhysChem properties of the nanomaterial as reported in the publication and the MIE, Key Event and/or Adverse Outcome Pathway reported for that nanomaterial. This exemplifies how the ERM identifier can be used to discuss nanomaterials in a FAIR way.

3.3.3. Gov4Nano

The Gov4Nano project (www.gov4nano.eu/) recently established the GO FAIR implementation network, AdvancedNano (www.go-fair.org/implementation-networks/overview/advancednano/), which will support implementation of the FAIR principles in nanosafety data management systems. To achieve the overarching aim to disseminate and collaborate on solutions for increasing FAIRness of nanosafety data, the project is working on a set of six case studies that demonstrate findability and accessibility (case study 1), interoperability (case study 2), and reusability (case studies 3 to 6) of nanosafety data. Case study 1 is focused on implementation of persistent identifiers and aims to develop strategies for the community to actively use such identifiers, including the ERM identifiers, focusing on barriers and incentives for uptake of solutions. Case study 2 focuses on the use of electronic laboratory notebooks to support increased implementation of workflows to standardize data formats and harmonize metadata annotation, including ERM identifiers, related to work developed in NanoCommons on the data life cycle and metadata requirements and needs at all stages of the nanosafety data lifecycle [28]. Case studies 3 to 6 aim to demonstrate reuse of omics, ecotoxicological, genotoxicity, and human occupational exposure data, respectively. Each case study will consider curation strategies involving implementation of ERM identifiers. For example, the omics case study is focused on providing reusable nanosafety-relevant omics data as input for tools and models in the Nano Risk Governance Portal, supporting the Nano Risk Governance Council developed in concert with the two other governance projects (RiskGONE and NanoRIGO). Reuse of omics data is highly dependent on integration with other types of data, such as physicochemical and exposure data. Implementation of the ERM identifiers supports this type of integration across diverse types of data and eventually enables interoperable reuse of relevant high-content omics data [2].

3.3.4 NanoSolveIT

The nanoPharos database (db.nanopharos.eu/), developed within the H2020 NanoSolveIT (www.nanosolveit.eu, [29]) and NanoCommons projects, is promoting accessibility, interoperability and reusability of curated datasets in a ready-for-modelling format. The nanoPharos database has been designed under the FAIR data principles to also include computationally derived data from simulations of nanomaterials at different levels of complexity. The database was further extended to include nanomaterials characterization data that can be enriched with a series of atomistic, molecular and structural descriptors. The database offers the possibility of including different batches and instances of the same nanomaterial, so as to monitor any physico-chemical transformations across its entire lifecycle. Thus, it is possible to use a single starting ERM and follow nanomaterial transformations/changes in different laboratories and/or projects and under different environmental conditions. The nanomaterials data can be linked with biological effects data to support complete *in silico* nanosafety evaluation. Currently, the NanoSolveIT project has registered 57 ERM identifiers to uniquely identify which nanomaterials the information in the database relates to. An additional 25 ERM identifiers were added to the nanoPharos database for a dataset derived from a recent publication [30]. Because this dataset was a subset of the original dataset, which was in NanoWiki 6 [31], the ERM identifiers were also added to NanoWiki.

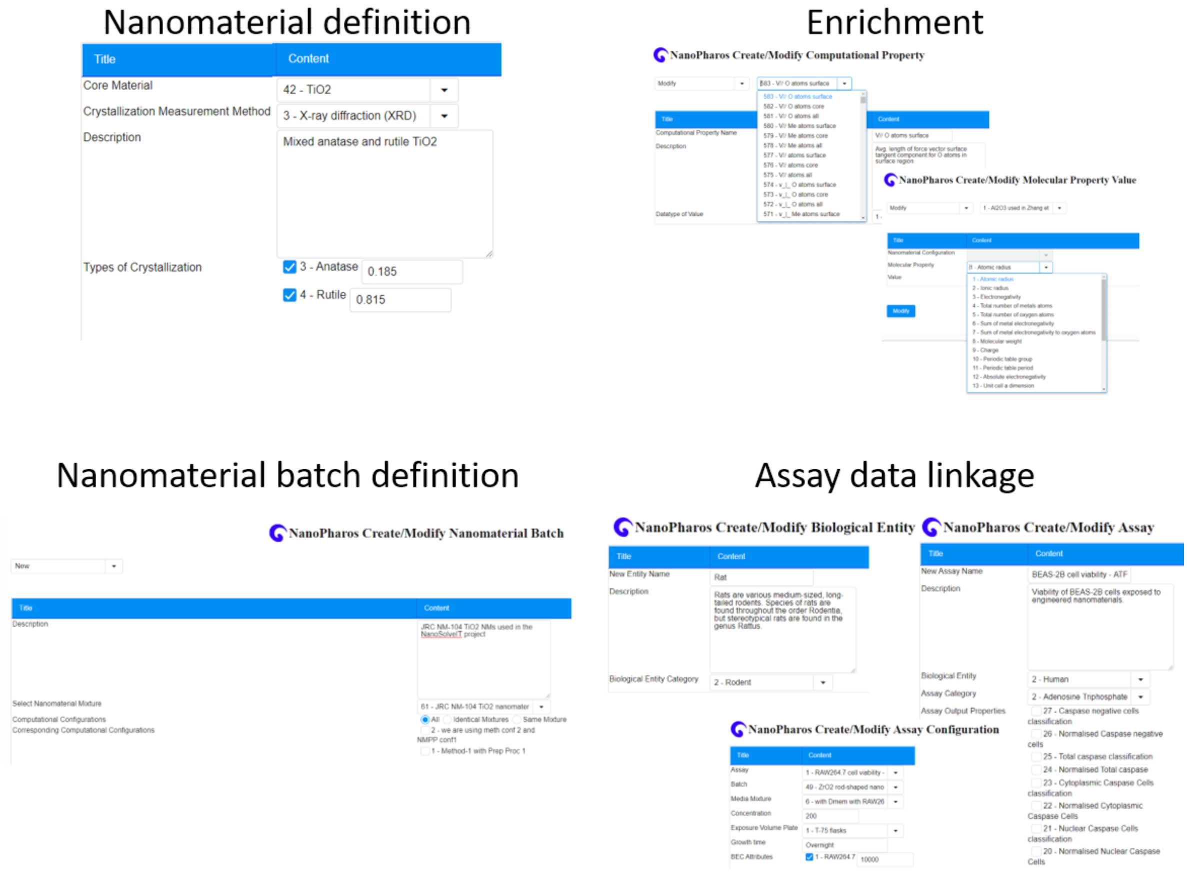

Figure 2. The nanoPharos database offers users the ability to define specific nanomaterials, enrich the main structure with molecular and atomistic descriptors, which can be linked with a specific ERM. The system offers the ability to define separate batches of the same nanomaterial, directly linked to the original material and to the subsequent biological assays. In this way, it is possible to track the lifecycle of the starting nanomaterial and its different batches using the ERM.

3.3.5 caLIBRAte

A total of nine ERM identifiers were used in the caLIBRAte project. The focus of the project was on specific NRCWE nanomaterials that already had their own identifiers. These NRCWE IDs were coupled to ERM identifiers in the caLIBRAte - eNanoMapper database at search.data.enanomapper.net/projects/calibrate. With this mapping (currently internal to the project), the ERM ids provide information on the chemistry, state and supplier of the nanomaterials.

3.3.6 SbD4Nano

The total number of preliminary ENMs for the SbD4Nano project is around 10. As each ENM is surface modified, an ERM identifier number will also be assigned for the new materials. The registered identifiers for the preliminary ENMs are labeled with batch number, date, chemical composition, supplier name, and code using eNanoMapper (search.data.enanomapper.net/projects/sbd4nano/). Within the project, toxicological profiles of ENMs are investigated *in vitro*. The eNanoMapper database [2] is also used for toxicological data management for the modified ENMs. This information is made available to other project partners through e-mail, databases, and as Markdown files. The properties of ENMs such as substance identification, substance and physicochemical characterization, matrix information, end-product information, sector applications, and process of synthesis from the supply chain participants, were registered on the shared project drive. Data generated by project partners is being entered using NanoSafety cluster Excel templates, with an integrated Materials datasheet, allowing easy ERM lookup during data entry, as described for the RiskGONE project database. While the data templates are shared between projects, the TemplateWizard populates the Materials datasheet with project-specific material identifiers.

For the SbD4Nano project, data extraction will be done analogously to the approach described for the RiskGONE project (see Section 3.3.2). Six ERM identifiers have already been used for this effort.

3.3.7 NanoTest

The FP7 NanoTest project (www.nanotest-fp7.eu) focused on ENMs used in nanomedicine. As the area of nanomedicine brings humans into direct contact with nanomaterials, data on ENM interaction with cells from eight different organs and tissues were studied. The NanoTEST project is completed and data is available to several ongoing projects (e.g. RiskGone) through the eNanoMapper Nanosafety Data Interface [2] https://search.data.enanomapper.net/. Six materials are linked to ERM identifiers, ERM00000068-ERM00000073. Genotoxicity, oxidative stress, cell viability, immunotoxicity and physicochemical characterisation data generated on blood, lung, brain, gastrointestinal tract, liver, kidney, vascular system and placenta models by the project partners can be retrieved using these identifiers.

4. Discussion

Seven nanosafety projects have begun using the ERM identifier, a great start. Clearly, not all questions on how the identifier will be used in practice have been answered as yet. Indeed, during the initial dissemination of the identifier, a number of questions arose. For example, do we really need another identifier? The Registry does not bar the creation of identifiers for nanomaterials that already have an identifier, however the purpose is primarily to create identifiers for materials that did not yet have a unique identifier.

A second question arose about the anonymity of the materials. As with the previous question, the Registry allows users to just register a label, however the Registry can host several additional types of optional information. The reason for this choice is that when we start research on a new material, we may not have experimentally established its identity, but we need to ensure we order and/or distribute the correct material. The new identifier serves that purpose.

A third issue is the likelihood that the same materials be registered more than once. This can happen when projects that operate independently and are not aware of each other, continue to conduct research on materials from a completed project. However, this problem is common to other identifier systems, and the equivalence of identifiers and how to handle it, has been studied [32,33]. A practical solution to this problem is to simply use an identifier mapping showing that two identifiers refer to exactly the same material. It is important to realize that this approach is a lot more tractable than accidentally using the same ERM identifier, and then later realizing that the materials are not identical. Indeed, there are plenty of solutions to handle the equivalence of identifiers, like the ELIXIR Recommended Interoperability Resource BridgeDb [32].

A fourth issue is that the information that is coupled to the identifiers is not always already publicly available. For some of the above use cases, that information is currently only available within the projects and there is currently no clear solution. Projects have to decide when and how to make this information available. Therefore, one should not use identifiers reserved by other projects and make assumptions as to which nanomaterials they refer to. Therefore, it is important to keep track of the material's authority, that is, who can positively identify the material. This can e.g. be a paper or a contact person in the project, but it needs to be linked to the identifier (something that is not consistently done as yet).

As more projects adopt the ERM, and as its utility and limitations are explored by the nanosafety community, it is planned that community consensus on appropriate uses of the ERM will be defined and documented, and solutions to the issues noted above reached collectively.

5. Conclusion

This communication describes the European Registry of Materials and the corresponding ERM as a persistent, unique identifier for nanomaterials and how various nanosafety projects have begun to use it. The identifier fills a need by allowing the projects to apply the FAIR principles to research output while it is being generated. The assignment of ERM identifiers already when selecting nanomaterials to test within a project facilitates the consistent identification of materials by multiple partners within the same project. Tools like eNanoMapper Template Wizard and an integrated ERM lookup on data entry alleviate the problem that ERMs are not easy to remember. It is important to note that how the ERM identifier is used is left to the user currently, and the use cases show examples of how the identifier is adopted. The full value will become clear when many more research outputs are publicly shared, as it will allow us to connect results accurately. The active uptake of the identifier in multiple data reuse case studies, and the intention of projects like NanoInformaTIX and HARMLESS to start using them, bodes well for the future of this persistent nanomaterial identifier.

Author Contributions: Conceptualization and methodology, E.W.; writing—original draft preparation, J.v.R., E.W.; writing—review and editing, J.v.R., A.A., M.C., M.D., T.E., N.J., E.M.L., I.L, G.M., P.N., A.G.P., D.A.W., H.Y, E.W.; project administration, J.v.R. All authors have read and agreed to the published version of the manuscript. The order of the authors is as follows: the first author and last author were fixed and the other authors are listed in alphabetical order of their last name.

Funding: This research was funded by European Union Horizon 2020 Programme (H2020) under grant agreement nº 731032 (NanoCommons), grant agreement nº 814425 (RiskGONE), grant agreement nº 814572 (NanoSolveIT), grant agreement nº 814401 (Gov4Nano), and grant agreement nº 862195 (SbD4Nano).

Acknowledgments: We thank GitHub for providing their open innovation platform.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kochev, N.; Jeliazkova, N.; Paskaleva, V.; Tancheva, G.; Iliev, L.; Ritchie, P.; Jeliazkov, V. Your Spreadsheets Can Be FAIR: A Tool and FAIRification Workflow for the ENanoMapper Database. *Nanomaterials* **2020**, *10*, 1908, doi:10.3390/nano10101908.
2. Jeliazkova, N.; Apostolova, M.D.; Andreoli, C.; Barone, F.; Barrick, A.; Battistelli, C.; Bossa, C.; Botea-Petcu, A.; Châtel, A.; De Angelis, I.; et al. Towards FAIR Nanosafety Data. *Nat. Nanotechnol.* **2021**, *16*, 644–654, doi:10.1038/s41565-021-00911-6.
3. European Commission *Cost-Benefit Analysis for FAIR Research Data - Cost of Not Having FAIR Research Data*; Brussels, 2018; ISBN 978-92-79-98886-8.
4. Wilkinson, M.D.; Dumontier, M.; Aalbersberg, Ij.J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.-W.; Santos, L.B. da S.; Bourne, P.E.; et al. The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Sci. Data* **2016**, *3*, sdata201618+, doi:10.1038/sdata.2016.18.
5. Jacobsen, A.; de Miranda Azevedo, R.; Juty, N.; Batista, D.; Coles, S.; Cornet, R.; Courtot, M.; Crosas, M.; Dumontier, M.; Evelo, C.T.; et al. FAIR Principles: Interpretations and Implementation Considerations. *Data Intell.* **2020**, *2*, 10–29, doi:10.1162/dint_r_00024.
6. Dittmar, P.G.; Stobaugh, R.E.; Watson, C.E. The Chemical Abstracts Service Chemical Registry System. I. General Design. *J. Chem. Inf. Model.* **1976**, *16*, 111–121, doi:10.1021/ci60006a016.
7. Heller, S.R.; McNaught, A.; Pletnev, I.; Stein, S.; Tchekhovskoi, D. InChI, the IUPAC International Chemical Identifier. *J. Cheminformatics* **2015**, *7*, 23+, doi:10.1186/s13321-015-0068-4.
8. Lynch, I.; Afantitis, A.; Exner, T.; Himly, M.; Lobaskin, V.; Doganis, P.; Maier, D.; Sanabria, N.; Papadiamantis, A.G.; Rybinska-Fryca, A.; et al. Can an InChI for Nano Address the Need for a Simplified Representation of Complex Nanomaterials across Experimental and Nanoinformatics Studies? *Nanomaterials* **2020**, *10*, 2493, doi:10.3390/nano10122493.
9. Karcher, S.; Willighagen, E.L.; Rumble, J.; Ehrhart, F.; Evelo, C.T.; Fritts, M.; Gaheen, S.; Harper, S.L.; Hoover, M.D.; Jeliazkova, N.; et al. Integration among Databases and Data Sets to Support Productive Nanotechnology: Challenges and Recommendations. *NanoImpact* **2017**, *9*, 85–101, doi:10.1016/j.impact.2017.11.002.
10. List of Pharmaceutical Compound Number Prefixes. *Wikipedia*. https://en.wikipedia.org/wiki/List_of_pharmaceutical_compound_number_prefixes
11. Maryanoff, B.E. First-Time Disclosures of Clinical Candidates. *ACS Med. Chem. Lett.* **2018**, *9*, 575–575, doi:10.1021/acsmedchemlett.8b00207.
12. *Markdown Tutorial*; https://www.markdowntutorial.com/
13. Carothers, G.; Prud'hommeaux, E. *RDF 1.1 Turtle*; W3C, 2014; https://www.w3.org/TR/turtle/

14. the FAIRsharing Community; Sansone, S.-A.; McQuilton, P.; Rocca-Serra, P.; Gonzalez-Beltran, A.; Izzo, M.; Lister, A.L.; Thurston, M. FAIRsharing as a Community Approach to Standards, Repositories and Policies. *Nat. Biotechnol.* **2019**, *37*, 358–367, doi:10.1038/s41587-019-0080-8.
15. Juty, N.; Le Novère, N.; Laibe, C. Identifiers.Org and MIRIAM Registry: Community Resources to Provide Persistent Identification. *Nucleic Acids Res.* **2012**, *40*, D580–D586, doi:10.1093/nar/gkr1097.
16. Wimalaratne, S.M.; Juty, N.; Kunze, J.; Janée, G.; McMurry, J.A.; Beard, N.; Jimenez, R.; Grethe, J.S.; Hermjakob, H.; Martone, M.E.; et al. Uniform Resolution of Compact Identifiers for Biomedical Data. *Sci. Data* **2018**, *5*, 180029, doi:10.1038/sdata.2018.29.
17. Hastings, J.; de Matos, P.; Dekker, A.; Ennis, M.; Harsha, B.; Kale, N.; Muthukrishnan, V.; Owen, G.; Turner, S.; Williams, M.; et al. The ChEBI Reference Database and Ontology for Biologically Relevant Chemistry: Enhancements for 2013. *Nucleic Acids Res.* **2013**, *41*, D456–D463, doi:10.1093/nar/gks1146.
18. Hastings, J.; Chepelev, L.; Willighagen, E.; Adams, N.; Steinbeck, C.; Dumontier, M. The Chemical Information Ontology: Provenance and Disambiguation for Chemical Data on the Biological Semantic Web. *PLoS ONE* **2011**, *6*, e25513, doi:10.1371/journal.pone.0025513.
19. Hastings, J.; Jeliazkova, N.; Owen, G.; Tsiliki, G.; Munteanu, C.R.; Steinbeck, C.; Willighagen, E. ENanoMapper: Harnessing Ontologies to Enable Data Integration for Nanomaterial Risk Assessment. *J. Biomed. Semant.* **2015**, *6*, doi:10.1186/s13326-015-0005-5.
20. Rocca-Serra, P.; Brandizi, M.; Maguire, E.; Sklyar, N.; Taylor, C.; Begley, K.; Field, D.; Harris, S.; Hide, W.; Hofmann, O.; et al. ISA Software Suite: Supporting Standards-Compliant Experimental Annotation and Enabling Curation at the Community Level. *Bioinformatics* **2010**, *26*, 2354–2356, doi:10.1093/bioinformatics/btq415.
21. Willighagen, E.; Najko Jahn; Nielsen, F.Å. The EU NanoSafety Cluster as Linked Data Visualized with Scholia. **2018**, 470931 Bytes, doi:10.6084/M9.FIGSHARE.6727931.
22. JKChang; Willighagen, E.; Winckers, L.; UanaF; GerganaTancheva123; Ehrhart, F. (Freddie); Rieswijk, L.; Hastings, J.; Jeliazkova, N.; Slaughter, W.; et al. *Enanomapper/Ontologies: Release 7 of the ENanoMapper Ontology*; Zenodo, 2021; doi:10.5281/ZENODO.4600986
23. Amos, J.D.; Tian, Y.; Zhang, Z.; Lowry, G.V.; Wiesner, M.R.; Hendren, C.O. The Nanoinformatics Knowledge Commons: Capturing Spatial and Temporal Nanomaterial Transformations in Diverse Systems. *NanoImpact* **2021**, 100331, doi:10.1016/j.impact.2021.100331.
24. Amos, J.; Hendren, C.; Papadiamantis, A.; Matzke, M.; Walker, L.; Svendsen, C. *NanoFASE Data Curation Manual - Manual and a Visual Guide to Curating Data*; H2020 NanoFASE project, 2020; https://infrastructure.nanocommons.eu/library/?category=Written+tutorial&audience=&organisation=DU&service=
25. European Commission. Joint Research Centre. *NANoREG Data Logging Templates for the Environmental, Health and Safety Assessment of Nanomaterials.*; Publications Office: LU, 2017; doi:10.2787/505397
26. European Commission. Joint Research Centre. *GRACIOUS Data Logging Templates for the Environmental, Health and Safety Assessment of Nanomaterials.*; Publications Office: LU, 2019; doi:10.2760/142959
27. Murugadoss, S. A Strategy towards the Generation of Testable Adverse Outcome Pathways for Nanomaterials. *ALTEX* **2021**, doi:10.14573/altex.2102191.
28. 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.; et al. Metadata Stewardship in Nanosafety Research: Community-Driven Organisation of Metadata Schemas to Support FAIR Nanoscience Data. *Nanomaterials* **2020**, *10*, 2033, doi:10.3390/nano10102033.
29. Afantitis, A.; Melagraki, G.; Isigonis, P.; Tsoumanis, A.; Danai Varsou, D.; Valsami-Jones, E.; Papadiamantis, A.; Ellis, L.-Jayne.A.; Sarimveis, H.; Doganis, P.; et al. NanoSolveIT Project: Driving Nanoinformatics Research to Develop Innovative and Integrated Tools for in Silico Nanosafety Assessment. *Comput. Struct. Biotechnol. J.* **2020**, S2001037019305112, doi:10.1016/j.csbj.2020.02.023.
30. Papadiamantis, A.G.; Jänes, J.; Voyiatzis, E.; Sikk, L.; Burk, J.; Burk, P.; Tsoumanis, A.; Ha, M.K.; Yoon, T.H.; Valsami-Jones, E.; et al. Predicting Cytotoxicity of Metal Oxide Nanoparticles Using Isalos Analytics Platform. *Nanomaterials* **2020**, *10*, 2017, doi:10.3390/nano10102017.
31. Willighagen, E. NanoWiki 6, Figshare, 2020; doi:10.6084/m9.figshare.1330208
32. van Iersel, M.P.; Pico, A.R.; Kelder, T.; Gao, J.; Ho, I.; Hanspers, K.; Conklin, B.R.; Evelo, C.T. The BridgeDb Framework: Standardized Access to Gene, Protein and Metabolite Identifier Mapping Services. *BMC Bioinformatics* **2010**, *11*, 5+, doi:10.1186/1471-2105-11-5.
33. Batchelor, C.; Brenninkmeijer, C.Y.A.; Chichester, C.; Davies, M.; Digles, D.; Dunlop, I.; Evelo, C.T.; Gaulton, A.; Goble, C.; Gray, A.J.G.; et al. Scientific Lenses to Support Multiple Views over Linked Chemistry Data. In *The Semantic Web – ISWC 2014*; Mika, P., Tudorache, T., Bernstein, A., Welty, C.,

Knoblock, C., Vrandečić, D., Groth, P., Noy, N., Janowicz, K., Goble, C., Eds.; Springer International Publishing: Cham, 2014; Vol. 8796, pp. 98–113 ISBN 978-3-319-11963-2.