The European Nanotechnology Community Informatics Platform: Bridging data and disciplinary gaps for industry and regulators

Grant Agreement No 731032

Deliverable Report 10.3

Deliverable	D10.3 Initial Version of NanoCommons Sustainability Plan
Work Package:	WP10: [NA4 – Integration & Sustainability]
Delivery date:	M12 (Actual: 27/ 04 /2019, Month 16)
Lead Beneficiary:	Douglas Connect (DC) now Edelweiss Connect (EwC)
Nature of Deliverable:	Report
Dissemination Level:	Public

Submitted by:	Barry Harry (DC, EwC), Iseult Lynch and Anastasios Papadiamantis (UoB) and Beatriz Alfaro Serrano (BNN)
Revised by:	Iseult Lynch (UoB) and Christine Hendren (Duke)
Approved by:	Iseult Lynch, University of Birmingham

This project has received funding from European Union Horizon 2020 Programme (H2020) under grant agreement nº 731032

Table of contents

List of Abbreviations

AB – Advisory Board
AOP – Adverse Outcome Pathway
CEINT – Center for the Environmental Implications of Nanotechnology (Duke University, USA)
CROs - Contract Research Organisations
CS – Case Study
DC – Douglas Connect
DoA – Description of Action
EC – European Commission
ECHA – European Chemicals Agency
EFSA – European Food Safety Authority
ELIXIR – A distributed European life sciences infrastructure for biological data
EMCC – European Material Characterization Council
EOSC – European Open Science Cloud
ERIC – European Research Infrastructure Consortium (s legal status)
ESFRI – European Strategy Forum on Research Infrastructures
EOSC – European Open Science Cloud
EU – European Union
EUON – European Union Observatory for Nanomaterials
EURL ECVAM – EU Reference Laboratory for alternatives to animal testing
EwC – Edelweiss Connect
IP – Intellectual Property
KPI – Key Performance Indicator
MERIL – Mapping of the European Research Infrastructure Landscape
NIA – Nanotechnology Industry Association
NICK – NanoInformatics Knowledge Common (from CEINT)
NM – Nanomaterial
NPs – Nanoparticles
NSC – NanoSafety Cluster
OECD – Organisation for Economic Cooperation and development
QA/QC – Quality Assurance and Quality Control (of data)
QMRF – QSAR Model Reporting Format
QSAR – Quantitative Structure Activity Relationship
REACH – Registration, Evaluation, Authorisation of Chemicals
RI – Research Infrastructure
RGC - Risk Governance Council
SMEs - Small and Medium Enterprises
SOPs - Standard Operating Procedures
WG – Working Group
WP – Work Package
WPMN – Working Party on Manufactured Nanomaterials (OECD)
VO – Virtual Organisation

Existing long-term infrastructure platforms

ELIXIR – An intergovernmental organisation that brings together life science resources from across Europe. These resources include databases, software tools, training materials, cloud storage and supercomputers. The goal of ELIXIR is to coordinate these resources so that they form a single infrastructure. This infrastructure makes it easier for scientists to find and share data, exchange knowledge and find the right tools and training materials to fit their needs.

MERIL – Mapping of the European Research Infrastructure Landscape: EU H2020 Project which provides a comprehensive database and inventory of the main research infrastructures (RIs) across all European countries and scientific domains, including the humanities and social sciences

EOSC Hub – the European Open Science Cloud Hub brings together multiple service providers to create the Hub: a single contact point for European researchers and innovators to discover, access, use and reuse a broad spectrum of resources for advanced data-driven research.

EUON – the European Union Observatory for Nanomaterials is a public facing arm of the European Chemicals Agency, aiming to provide information about existing nanomaterials on the EU market.

EURL ECVAM – the EU Reference Laboratory for alternatives to animal testing is an integral part of the Joint Research Centre (JRC), the science and knowledge service of the EC. The mandate of EURL ECVAM is specified in EU legislation on the protection of animals used for scientific purposes and includes a number of duties to advance the Replacement, Reduction and Refinement (the Three Rs) of animal procedures. *In silico* approaches provide an important input towards this goal, and ECVAM undertakes validation of models, as well as validation of new *in vitro* experimental approaches.

ERIC – the European Research Infrastructure Consortium, is a specific legal form that facilitates the establishment and operation of Research Infrastructures with European interest on a non-economic basis. The range of ERICs is broad, and there are currently 20 in existence, as per the ERIC Landscape[1].

ESFRI – the European Strategy Forum on Research Infrastructures, is a strategic instrument to develop the scientific integration of Europe and to strengthen its international outreach. It publishes a roadmap every 2 years.

[1] https://ec.europa.eu/info/research-and-innovation/strategy/european-research-infrastructures/eric/eric-landscape_en

1. Summary

During the first year of the project a strategy was developed for ensuring sustainability of the NanoCommons research infrastructure and including the tools and services developed during and beyond the current project funding lifetime. A key aspect of this ongoing activity is identification of the market needs and business planning to address these needs. A process was put in place to develop the sustainable business plans, including a timescale for undertaking the supporting conversations with the market.

An initial sustainability plan was established including a range of options that will be analysed further and road-tested with potential users, and a process has been outlined to develop these plans further and test their feasibility through community and market engagement. The second year of the project will focus on further fleshing out and road-testing the various aspects of the business plan. In particular, industry case studies and use cases of industrial and business relevance are being identified and elaborated, so as to identify matches between real industrial needs and solutions that can be developed and provided within the NanoCommons framework and context.

Case studies were selected as the most useful way to demonstrate what NanoCommons can do and what support it can offer users, and as part of this activity, a key component will be to assess via feedback from three users what the value-add for them was in the short, medium and longer terms, and how this translates into economic benefit. The NanoCommons sustainability team will then evaluate the cost to the NanoCommons partners of providing the services and develop appropriate pricing and cost models for providing these services on a paid basis. Nine case studies are currently underway (Section 4), in collaboration with academic, consultancy, regulatory (national and European), industry and supra-national organisations, to span the full breadth of potential users and customers for NanoCommons as identified in Section 3.

Leveraging of experiences and best practice from across the infrastructure and e-infrastructure sector will also be key to long term success of NanoCommons. Indeed, integration with, and co-service approaches with other e-infrastructures have been identified as an essential step in sustainability, and our next phase will also focus increasingly on considering the value-add our services can offer to other communities and other e-infrastructures, and what we can leverage from them. For example, closer integration into the European Open Science Cloud (EOSC) and with the European Strategy Forum on Research Infrastructures (ESFRI) will be actively sought over the remaining duration of the project to ensure greater visibility and integration with ongoing large-scale activities.

Additionally, close cooperation with other projects involved in modelling of nanosafety is also being built in, including with the recently funded nanoinformatics projects NanoSolveIT and NanoInformaTIX, and the risk governance projects Gov4nano and RiskGone. Ongoing collaboration with OpenRiskNet (and eNanoMapper) is also be leveraged to co-develop strategies to ensure the sustainability of the generated outputs of the different projects.

2. Introduction

WP10 of NanoCommons was designed to tackle the issues of integration (of currently disparate tools, datasets and approaches) and sustainability of the research infrastructure built up over the 4 years of the project. By having this WP run throughout the duration of the project, the goal is that consideration of long-term sustainability becomes embedded into everything that the project does, such that when the initial project funding ends there is a durable and sustainable infrastructure and business model to maintain the platform as developed, and ideally to continue its development and improvement into the future.

NanoCommons is designed and set-up to function as a long-term community infrastructure for nanosafety nanoinformatics, providing tools and expertise to researchers, industry and regulators to support the advancement of *in silico* nanosafety science. Core to this is harmonisation of datasets, and the integration, benchmarking and documentation of current nanoinformatics tools, and development of a business model that allows them to be maintained and continuously updated beyond the project funding duration. Hence, in the first instance, User needs (academic, industry and regulatory) in terms of access to data, modelling tools, and nanoinformatics expertise, and the potential role of NanoCommons as a marketplace for nanosafety nanoinformatics tools and services will be assessed. This User-derived information will provide the evaluation ground for the selection and prioritisation of nanoinformatics tools for integration into NanoCommons, and will help in the prioritisation of documentation and benchmarking activities to facilitate the uptake of the tools and services by industry and regulators. Additionally, insights gained from the provision of TA regarding where the community needs and priorities lie, and from the Use case studies (WP9) designed to showcase the capabilities of the NanoCommons platform and tools integration, will further drive the development of the services that can be offered to Users. Insights regarding the barriers experienced by users in terms of their utilisation of nanoinformatics and *in silico* nanosafety approaches in real-life scenarios, gained during the first 3 years of the project, will further focus and drive NanoCommons's efforts towards full integration of computational modelling tools into industry and regulatory workflows. The development of customer-focussed solutions will establish the NanoCommons KnowledgeBase as a central resource for *in silico* nanosafety research and decision making.

NanoCommons is not the first project to face these issues, and indeed sustainability of outputs from all EU-funded nanosafety projects is a well-recognised issue, and has been the topic of a Task Force within the Nanosafety Cluster. Task 10.2 will align with ongoing activities to develop the access (to services and expertise) and business models, including providing a platform for subsequent EU-funded projects and/or allowing privileged pay-to-use access to premium services for a fee, for example. Sustainability beyond the project will be sought by developing both public and business components to the infrastructure while maintaining its core resources as open access ones. The goal of this deliverable is to outline the current thinking and progress towards the development of short, medium and long-term (2, 5 and 10-year) sustainability plans for NanoCommons.

Key customers for NanoCommons activities include the newly funded nanoinformatics projects (NMBP13) and the Risk Governance projects (NMBP14), and forthcoming projects on Safe by Design nanomaterials (NMBP15), as well as the Malta Initiative projects where NanoCommons will provide electronic notebooks and templates for the Round Robins. Longer term customers include the EU Observatory for Nanomaterials (EUON) which has a remit to increase the accessibility of data relating to the safety and risk assessment of nanomaterials for the wider public. Thus, part of the sustainability plan is the sharing of the NanoCommons Data Management Plan to support other projects to align with our approaches and thus enhance the overall quality, completeness and FAIRness of nanosafety data. Part of this work will also include development of a costing model to allow forthcoming projects, be they national, EU or individual person (e.g. Marie Curie Postdoctoral Fellowships) to include some of the NanoCommons data management and nanoinformatics services or expertise into their projects to cover the data management and/or nanoinformatics tasks.

To place the NanoCommons resources on a sustainable footing for the longer term, a business plan for securing future access provision (i.e. to support users in their accessing of the tools and services) and iterative development of the overall platform and set of associated infrastructure facilities (e.g. through continued research and development, as well as continued integration of tools and services developed and offered by others) will be developed. The business plan will include identification of industrial needs and problems, and the matching solutions from within the scope of current and near-term capability of NanoCommons, as well as developing a roadmap of activities to address needs articulated by potential users that NanoCommons cannot yet address for which dedicated funding can be sought. Additionally, the implementation of the sustainable concept will be supported by testing of this business plan during the runtime of the project. Further shaping and modifications according to the feedback gained during the sustainability-testing-phase is envisaged which will lead to a sustainable integration of the NanoCommons infrastructure.

3. Identification of key customer groups

3.1 Supra-national organisations addressing nanosafety

OECD

The Organisation for Economic Co-operation and Development (OECD) works with governments to understand what drives economic, social and environmental change, including supporting the development of standardised and harmonised testing methods to support the mutual recognition of regulatory approval worldwide, thus reducing costs for business and enabling global trade. As part of the Malta initiative projects, to accelerate the revision of the chemical testing guidelines for use with nanomaterials, NanoCommons is providing infrastructure support including access to the electronic notebooks and integrated standard operating procedures (SOPs) and data collection templates annotated to allow direct integration into the database, and indeed has already shown how this case works via a round-robin use case within the ACEnano H2020 project. As the approaches are not nanosafety-specific, such approaches and tools could also be applied more widely in the work of the OECD, specifically that of the Working Party on Manufactured

Nanomaterials (WPMN) and other activities related to new and emerging technologies. During the next phase of development of the NanoCommons Sustainability plan, more detailed consideration of the services that NanoCommons can provide to the WPMN will be undertaken.

EUON

The European Union Observatory for Nanomaterials (EUON) aims to provide objective and reliable information on the innovation and safety aspects of nanomaterials on the EU market in order to collect and analyse information from a wide variety of publicly available sources, to complement existing information with external studies and to present information on uses and safety of nanomaterials in a layman's language. NanoCommons is working directly with EUON to integrate the NanoCommons Knowledge Infrastructure into EUON, and EUON will promote the NanoCommons tools and services that are useful to them.

The potential services that NanoCommons can provide to EUON and its host ECHA (the European Chemicals Agency) include predictive tools and models that utilise the NanoCommons KnowledgeBase to evidence when a NM counts as a unique nanoform (ECHA Nanoform) from an exposure and/or toxicological viewpoint, as well as enabling evaluation of the impact of small protocol changes or different instruments to measure the same end-point on the variability in the dataset, which will be important for understanding the degree of inherent variability from the methods. Additionally, the push by NanoCommons to capture batch-related information as well as NM-specific persistent identifiers will be an important step forward in facilitating integration of datasets from different sources. However, it is not yet clear whether EUON itself will pay for services from NanoCommons, or whether its endorsement of specific NanoCommons tools and datasets will encourage industry users to pay for NanoCommons's services. These are questions that will be teased out as part of the next phase of development of the NanoCommons sustainability planning.

EC4SafeNano: The European Centre for Risk Management and Safe Innovation in Nanomaterials & Nanotechnologies

An EU H2020-funded project, EC4SafeNano links together national centres for nanomaterials risk assessment, to provide a European Centre for Risk Management and Safe Innovation of nanomaterials. Over the last 3 years, EC4SafeNano has investigated the range of services that it could supply to stakeholders, and the business models available to it. Via numerous surveys, it has compiled an inventory of service needs and service providers, and is considering how it can act as a broker for service provision. Other issues that it is grappling with currently are how to keep its services up to date, and how to access the data and models needed to provide bespoke services. Thus, potentially the EC4SafeNano centre could be a customer for NanoCommons data and tools.

Among the services that EC4SafeNano is proposing is a platform to provide harmonized services and EU Certification Marks of Harmonized Services. This would involve defining and developing harmonized services, creation and promotion of EU Certification Mark(s), and provision of training and auditing services related to the harmonised services. The business proposition is that the centre would collectively agree expertise and EU Certification Mark(s). A Harmonized or an EU Certification Mark(s) could also be developed based on an internal demand for strategic purpose (a new market). Earning can rely on customer funds asking and paying for the development of a new Harmonized

Service, or a new EU Certification Mark for a Harmonized Service. Earning can rely also on Service providers asking to be trained to be able to provide Harmonized Service. Earning can also come from Service Provider paying for using an EU Certification Mark(s).

European Nanorisk Governance Council

The three projects funded under NMBP-14, and which started in 1st January 2019, have as a key outcome the establishment of a transparent, self-sustaining Risk Governance Council (RGC) with representatives from EU stakeholders, member states, industry and civil society. The RGC will act as a science-based governance body to ensure responsible innovation in NMs and nanotechnologies, nanosafety and to provide responsible two-way communication with stakeholders and civil society, based on high quality information. Beyond these project durations (2019-2022) the RGC should be self-sustaining, financially and scientifically. Thus, the RGC will need access to up-to-date data, and state of the art predictive models capable of addressing the needs of future NMs and future products, as well as future regulatory scenarios. NanoCommons is already collaborating with these projects to ensure that its platform and tools are embedded from the outset into the RGC framework, as part of the provision of scientific decision making capability. Definition of specific services that can be provided to the RGC, and cost models for these services, will be worked out over the subsequent iterations of the NanoCommons business plan.

European Network for Pilot Production Facilities and Innovation Hubs (EPP)

The European Network for Pilot Production Facilities and Innovation Hubs[2] is a Coordination and Support action providing a market place like solution for producers and projects involved in the scale-up of production of nanomaterials or nano-enabled products. Its mission is to accelerate the development of technology-enabled products, and its main "product" is an online tool that connects the European Network of Pilot Production Facilities with SMEs, Start-ups and Large enterprises in order to help overcome the so-called valley of death between research and commercialisation. Among its activities currently being initiated is a working group on safety, which is particularly relevant given that these pilot line projects are typically addressing the scale-up of production of ENMs and/or nano-enabled products. This working group (WG) will address issues such as safety, modelling, characterisation, and standards. The main activities of this WG would be mapping needs and developing collaboration related to sharing of facilities, characterisation equipment; safety, standards and modelling. Several of these activities map directly to aspects of NanoCommons, such that the EPP and/or some of its constituent projects would also make an excellent case study for development within NanoCommons, to scope out the sorts of services that NanoCommons could provide to pilot line projects individually or collectively via the EPP.

The EPP also have a WG on market and finance intelligence, whose goal is to collect and provide commercial and market information and guidance to the pilot production operations and user cases. Among the objectives listed for this WG are to: support business planning, development and sustainability; support definition of the most appropriate business model, including open access opportunities; profile customer and market sector and develop an activate market place(s) and

[2] https://www.eppnetwork.com/

facilitate brokerage opportunities. Thus, there are also opportunities for NanoCommons WP10 partners to collaborate with the EPP on development of mutually supportive sustainability options and opportunities.

3.2 Industry users

The NanoCommons project will work to provide the necessary resources to relevant industry stakeholders, which will decrease the burden of data collection for product developers and industrial users. The tools and services offered through NanoCommons, in combination with the project's commitment to FAIR and Open data, can promote data innovation and lead to positive socioeconomic impact. This will be achieved through the facilitation of further data exploitation and translation of the results to new products and services and safe-by-design approaches opening up avenues for tools to be applied into new and emerging industrial sectors and regulatory settings. The value of FAIR data has been strengthened following the publication, in March 2018, of a report commissioned by the Danish Agency for Science and Higher Education titled "Preliminary analysis: introduction of FAIR data in Denmark" (FAIR). The main scenario of the report was based on the most conservative calculation assumptions found in the literature and disregarded the benefits that would fall to other countries if the FAIR data concept is introduced in Denmark; nor have any benefits gained from a FAIR data collaboration with other countries been included in the calculations. The results stated that the net present positive socioeconomic value, if it was assumed that 50 per cent of all research data in Denmark will comply with the FAIR principles, would add up to about DKK 2 billion (€268 million) over a 40-year period.

Among the tools and approaches proposed by NanoCommons to support users in marking their data FAIR and Open, and increase the utility and intrinsic value of the data itself are:

- Enhanced data quality through implementation of Quality Assurance / Quality Control (QA/QC) steps and processes from the point of experimental design and data generation, rather than managing the data at the end of projects as is often the case currently;
- Workflows for annotating and uploading datasets and for integrating databases, thereby reducing data-loss and embedding knowledge management practice in the next generation of nanosafety researchers;
- Enhanced usefulness of the datasets generated, as a result of their associated metadata, SOPs, ontology mapping, and accessibility (Open and useful download formats) both for modelling and regulatory decision-making purposes, as a result of the QA/QC approaches indicated above;
- Documentation of NanoCommons (and external partner) tools and their relative advantages and disadvantages, and data requirements, compliant with the guidelines of EURL ECVAM, such as utilising the reporting form (QMRF) templates for QSARs (Lamon), and the emerging EMMC model reporting templates for materials models, thus enhancing their utility for use in NMs regulatory dossiers or in building weight of evidence;
- Workflows and online notebooks to support the running of nanosafety testing, including the underpinning SOPs that will support the hand-over of routine nanosafety assessment from research labs to Contract Research Organisations (CROs), supporting the transition to the standard regulatory model whereby EHS aspects are out-sourced to specialist CROs.

NanoCommmons will facilitate industry to develop products with safety(-by-design), recyclability and functionality in mind, and provide tools to allow industry to vary their parameter space and predict optimal properties for safety/recyclability and functionality from the product design stage. It is expected that the NanoCommons tools will enable movement of nano-containing products from conceptual safer designs to commercial equivalents where the modelling tools will allow industry to vary their parameter space and predict optimal properties for safety/recyclability and functionality from the product design stage. While the tools will be suitable for all types of industry, the initial demonstrations and target case studies will be with small and medium enterprises (SMEs) and will be targeted towards NM manufacturers initially, where many SMEs operate. Tools for NM formulators / product developers will also be developed, but these are expected as later demonstrations, and/or through integration of existing tools such as GUIDEnano.

Thus, NanoCommons is initially targeting SME NM producers who will be interested in the Safety-by-Design approaches and the access to tools and model, CROs involved in nanosafety testing and larger companies with several NM variants around a common core chemistry, who will be interested in the tools for distinguishing nanoforms and predicting their behaviour.

3.3 Academic users

The increasingly data intensive character of scientific research and the need for high quality data is a key driver for NanoCommons to integrate disparate (and vulnerable to loss) research data, tools and models as a key step to securing streamlined access to data management tools for the nanosafety research community. Thus, NanoCommons tools and services target a range of academic users including:

- Experimentalists sitting on existing datasets, who will be encouraged and trained to integrate their datasets into existing repositories or the NanoCommons knowledgebase;
- Experimentalists in currently running projects / future data generators in order to embed data management processes early and thus prevent the continued generation of disparate datasets that are not suitable for automated integration into existing data repositories;
- Modellers who want to access the existing datasets as they become integrated into NanoCommons and who wish to integrate their tools into the NanoCommons platform to enable automatic access to data and to allow others to utilise their tools;
- Nanosafety cluster, Characterisation Cluster and Materials Modelling cluster projects as users of the NanoCommons data, models, tools and ontologies thus driving further integration of these communities and bootstrapping of progress and impact.

The need for high quality data will be addressed through the establishment of clear data quality concepts to support generation of highest quality data suitable for modelling and regulatory risk assessment and decision making and to allow the consolidation of such data within common open access community data holdings.

The optimisation of new integrated models and tools to enable the further mining of integrated and highly dimensional data-sets, functional analyses to support key concepts such as grouping and read across and determination of risks from NM exposures and the associated uncertainties, will also facilitate the experimental and modelling process. These will be linked with clear guidelines for the integration of data capture and data management approaches into experimental workflows to both enhance data quality and facilitate simplified data transfer to public repositories to ensure future data availability and sustainability.

The enhanced interoperability and accessibility of data, enabled by NanoCommons will thus provide a strong basis for development of approaches for innovation governance in nanotechnology, including the development of intelligent testing strategies, support for nanotechnology product development, and increasing public engagement in nano-safety research and risk governance.

4.0 Leveraging best practice from other e-infrastructures and KnowledgeBases

As noted in the introduction, NanoCommons is not the first project, nor will we be the last, to grapple with issues of sustainability. Europe has a long tradition of research infrastructure projects, and a growing track record of success in terms of e-infrastructures and their long-term maintenance, from which NanoCommons can learn and take inspiration. The UniProt database undertook an analysis of the various models for sustainability, which was published open access, and from which a number of approaches of relevance to NanoCommons can be abstracted (Gabella). These are shown schematically in Figure 1 and are grouped into three categories depending on who is paying for the services: Public funding (including via academic users), Commercial Users and Third party. These are described in outline next, and whether or not they are relevant to NanoCommons will be considered in the outline Sustainability plan.

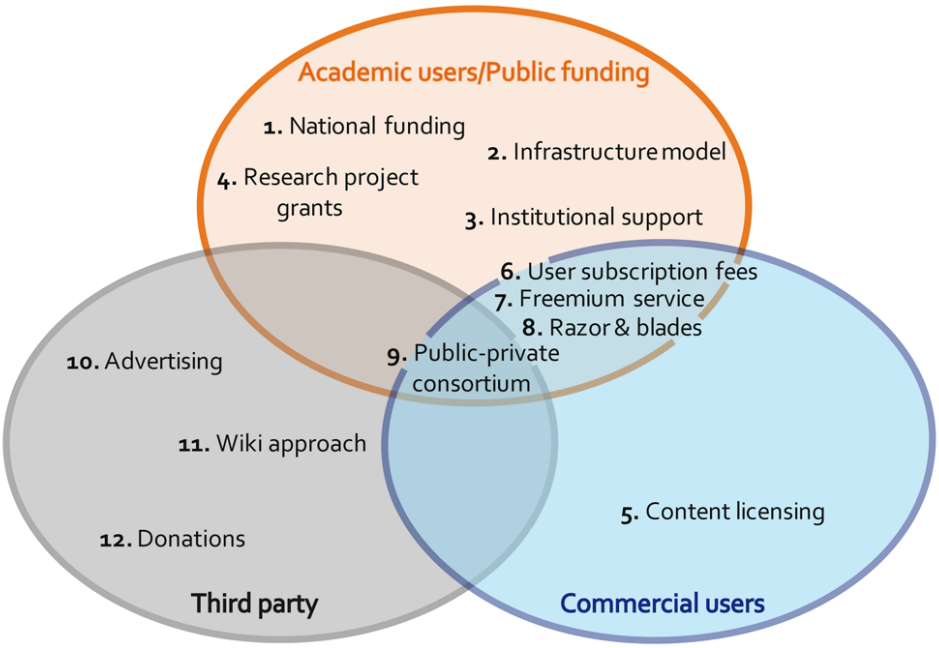

Figure 1: Funding Models for life sciences infrastructures and knowledgebases. From Gabella C, Durinx C and Appel R. Funding knowledgebases: Towards a sustainable funding model for the UniProt use case https://www.ncbi.nlm.nih.gov/pubmed/29333230

1. **National funding:** governmental agencies fund the infrastructure directly, through non-cyclical funding programmes. For research infrastructures, funds derive directly from national research budgets, and users may be charged for some subscriptions or special services.

2. **Infrastructure model**: funding agencies pay directly for data resources as a necessary part of the research infrastructure, through a percentage of the research funding that is specifically set aside. The grants themselves are only allocated to research projects. A percentage of each grant is then retained and assigned to a budget for data stewardship, and subsequently redistributed among the relevant infrastructures, including knowledgebases (Ember). The funding agencies can be national or private and contribute financially in proportion to the grant volume they allocate to research. This model is not implemented yet as a funding model for life science knowledgebases.

3. **Institutional support:** universities or institutions have their own repository/data bank that is maintained through the "internal" institutional funds. Grants can be cyclic or long-term, and usage may be restricted to the institution's members or be open to the worldwide community.

4. **Research project grants:** competitive cyclic research or dedicated resource grants from funding agencies. For EC research infrastructures, this can be infrastructure project funding, directly or via mapping to the ESFRI roadmap[3], or subsequent establishment of an ERIC. Access is free for the user as it is paid by the project (Poux). Most of the databases and knowledgebases in the life sciences are supported by this type of grants.

[3] https://www.esfri.eu/roadmap-2018

5. **Content licensing/industrial support model**: requires commercial users to pay a fee for access to the data and for-profit reuse, whereas data are free for non-commercial users (Maron).

6. **User subscription fees:** users are charged on a time basis (e.g. every month or year) or on download sizes, and they have access to the entire database. At the end of the validity, the subscription must be renewed to continue the access.

7. **Value-added/asymmetrical pricing model (freemium service)**: a basic data set within the database is freely available to anyone. Individual scientists or companies that are willing to pay a higher fee can buy additional levels of service, better data access or additional tools and resources (Bastow).

8. **Infrastructural razor & blades**: an attractive, inexpensive or free initial offer ("razor") encourages continuing future purchases of follow-up items or services ("blades"). (Ferro)

9. **Public-private consortium:** a mixture of funding from government bodies and industries. The funders mandate the research subjects and supporting companies do not receive priority access.

10. **Online advertising and corporate sponsorship:** corporate sponsorship is part advertising and part deal making —the corporation pays to support a database that provides value to its potential customers.

11. **Open source volunteering** (or **wiki approach**): replacing part of data curation by community participation can be attractive as it has a low cost (Gabella 2). It depends, however, on drawing contributions from busy users. In addition, contributions tend to be sporadic, leaving many gaps. Hence it can only replace (a small) part of curation and therefore still requires funding for curation, software engineers, storage space, and operating costs. There are also challenges in usability, authorship recognition, and information reliability (Chen)

12. **Donations**: philanthropic funding such as grants and donations can generate income. They partly depend on the impact and awareness of the (user) population.

In reality, most knowledgebases rely on **diversified** multiple funding streams, a so-called **Mixed models' approach**. This has the obvious advantage of increasing resilience if one source disappears. The models were summarized in the original report for comparison (Gabella), as shown in Table 1, comparing its compatibility with open access policies, its equity among the potential users and institutions, i.e. whether or not wealthier institutions or certain users are particularly favoured, the forecasted stability of the model over time, and the key dependency of each funding stream. The dependency of the funding is crucial to describe the vulnerability of the models and also needs to be taken into account when setting up a mixed model. The best funding model would combine models that are dependent on different factors.

Table 1: Summary of the characteristics of the different Funding Models for life science infrastructures and knowledgebases

#	Name of the model	Compatible with open access?	Potential for equity of users or institutions	Stability forecasted over time	Key dependency
1	**National funding**	**Yes**	**High**	**Stable**	National economic situation
2	**Infrastructure model**	**Yes**	**High**	**Stable**	Research spending by funding agencies
3	**Institutional support**	**Yes**	**High**	**Stable** or Cyclic	Institutional funds availability
4	**Research project grants**	**Yes**	**High**	Cyclic - grants renew every 3–5 years	Infrastructure/research spending by funding agencies (EU funding)
5	**Content licensing /industrial support model**	No	Low	Function of usage	Commercial partner
6	**User subscription fees**	No	Low	Function of usage	Usage
7	**Value-added/asymmetrical pricing model (or freemium service)**	Not completely	Low	Function of usage	Usage
8	**Infrastructural razor & blades**	No	Low	Function of usage	Usage
9	**Public-private consortium**	**Yes**	**High**	**Potentially stable**	Commercial partner
10	**Online advertising & Corporate sponsorship**	**Yes**	**High**	Function of usage	Usage, commercial partners
11	**Open source volunteer (wiki approach)**	**Yes**	**High**	Highly dependent on participation	Willingness to contribute
12	**Donations**	**Yes**	**High**	**Potentially stable**	Partners

The European Strategy Forum on Research Infrastructures (ESFRI) roadmap 2018 advises urgent establishment of a convergent policy of funding mechanisms for e-Infrastructures at the various levels (institutional, regional, national, European) (ESFRI Roadmap). Such policy could include support and financing of e-Infrastructures for scientific users, providing incentives to researchers to generate FAIR and reproducible data, as well as the development of enabling e-tools/e-technologies and the mainstreaming of support actions addressing e-needs of all levels of intervention.

5.0 NanoCommons Sustainability Plan – Initial Version

5.1 Objectives

Our objective is to make the NanoCommons platform attractive for a community of users by providing solutions that solve important problems and satisfy the informatics needs of the nanomaterials safety community. An example of such a need is that SMEs have difficulties understanding how they respond to the emerging complex science and regulations associated with nanosafety. NanoCommons should strive to develop services and tools that address such issues and will then have customers. Thus, engagement with potential users and customers of the NanoCommons Tools and Services and indeed co-development of solutions where appropriate, must be a central aspect of the project's implementation. In the absence of such engagement with the customer pool, NanoCommons will develop as a concept and set of resources without real committed users and little realistic chance of sustainability beyond the existing project funding.
NanoCommons, as a starting e-infrastructure, must excel and exceed the usual expectations and deliverables of research consortia projects which deliver meetings and reports with little actionable higher quality service development. A key focus for the ***project is to offer a toolbox of services and tools*** that SMEs need and can actually use and are willing to pay for.

We will define the following elements in developing a NanoCommons sustainability plan:

1. Definition of the platform aims;
2. Definition of its content and structure;
3. Definition of access and/or business models for/to the platform;
4. Definition of the contributors to the platform;
5. Definition of the users of the platform;
6. Definition of the services and tools provided by the platform;
7. Definition of trainings: teaching the contributors and users to use the platform and its content in a proper way;
8. Support service for users and contributors of the platform.

We anticipate from early market feedback the following business services as targets for NanoCommons:

- **Assessment expertise service -** provide consulting services supporting addressing of regulatory concerns;
- **Setup assessment for a particular concern** - e.g. assessment for a chemical substance; come up with a standard technical assessment; assess if addition for the nano-size related assessments is required;
- **Self-assessment support tool**: enabling companies to do basic assessment work by themselves.

The information assembled in the NanoCommons data infrastructure should take the end use of the above services into account. Which data, metadata, parameters is required to support the services and tools? If data gaps are identified in an assessment which modelling tools can be used to fill the gaps? If uncertainty in modelling predictions is too high, which experimental method should be used

to adequately address the data generation need? How will a quote for such a service be prepared, negotiated and delivered?

We anticipate development of the following components to the structure of the NanoCommons Platform supporting such service provision:

a) Provision of a standard technical assessment – combination of expertise and supporting resources and tools;
b) Providing assessments determining the information and evaluation extensions required for nanoforms;
c) Providing testing strategies supporting tiered approaches to assessment;
d) Identify which test results can be used from the NanoCommons database;
e) Which additional tests need to be done and how to feed their data into the NanoCommons platform and applications.

We will consider the following components of the business model with regards to access:

- Open Access to core reference resources will be supported with free access provided to users;
- Open Access fees with regards to quality data publication accompanying open access publication – disruption of the current stagnant publishing monopoly is called for;
- "Testing strategy" for fees – data generation will be quoted and charged for;
- Incentives model to feed new data into the NanoCommons Platform – provide value to end user owning data e.g., availability of analysis and insight into their own data once uploaded;
- Providing premium services to professional users requiring a fee.

Contributors of data to the NanoCommons Platform will be required to fulfil quality criteria with regards to description of the datasets and protocols loaded. Advice on measures to respond to failure to meet quality criteria will be supported by the NanoCommons helpdesk. The criteria will include adherence to emerging best practices associated with the FAIR principles including ethics and privacy guidelines addressing data access conditions. Procedures will be established for both raw and processed data.

Users of the NanoCommons Platform will include:

- Small companies – producing or incorporating nanomaterials in their products and solutions;
- Larger companies – producing or incorporating nanomaterials in their products and solutions;
- Regulators – evaluating regulatory issues associated with nanomaterials in products;
- Policy makers – evaluating policy issues associated with nanomaterials in products;
- Researchers and Scientists – supporting the use of data and tools in research activities.

5.2 Case Studies Demonstrating NanoCommons services for customers

Within the project, and specifically in WP9, different Case Studies are being defined and implemented as a means to demonstrate the capabilities of the NanoCommons e-infrastructure tools and services. These Case Studies are, in the first instance, intended as a demonstration of the Tools and Services which will be offered by the NanoCommons Platform, to exemplify the sorts of projects, challenges or "needs" that stakeholders can bring to NanoCommons in order for us to develop solutions – in some cases the solutions may already exist but have not yet been made available in a user friendly format, while in other cases a bespoke solution may need to be developed. During the project lifetime this development can potentially (subject to selection by the User Selection Panel for TA support) be developed via the project's TA provision, but in the longer term will be the sort of bespoke services for which NanoCommons partners can charge.

Parallel to the Case Studies defined in WP9, additional Cases Studies have been defined within WP10 under Task 10.2 (Sustainability Plan), in order to identify potential users of the NanoCommons e-infrastructure platform, and to showcase how NanoCommons can solve their needs. The goal with these case studies is that the clients / users will then provide testimonials as to the value that NanoCommons represented for them and will act as ambassadors for NanoCommons. Thus, the WP10 Sustainability Case Studies are intended to provide the users with examples of specific services that can later be rolled out with a cost model. In this case, NanoCommons will actively engage with these pre-identified potential users, identify problems they have that can be solved by NanoCommons, and WP10 will support the infrastructure development work on each of the specific user Case Study projects to provide solutions to the Users' needs (see Table 2). The goal is to develop these conversations with the market during Q1-Q2 2019, determine problems to be solved, and provide a matching and prioritisation where we can provide solutions from NanoCommons, within Year 2 (2019) and later (2020-2021).

As a main goal of NanoCommons is to create a sustainable e-infrastructure platform, besides the lifetime of the project, the users/clients of the platform will count on extra-cost for the services/tools offered by NanoCommons. These identified cases studies will help us to screen the market and to be able to generate a business plan for NanoCommons, based on the gained experiences, evaluating the value of the offered services and the estimation of costs:

- How to identify potential users, how to contact them and how to gain their confidence;
- How to identify the needs of the user;
- Expected service by the user, once the needs are identified (versus service offered by NanoCommons);
- Estimation of Costs: resources needed for the implementation of the service (time invested in the scoping of case study and implementation, personnel costs);
- Added-value of NanoCommons services;
- How to make the services more valuable to the users;
- Are users interested in paying for the services and what?

Table 2. List of Cases Studies (CS) identified so far within Task T10.2.

Nr.	Company	Category	Contact Partner (NanoCommons)	Identified Need	Kind of Case Study	Status of Case Study
1	INERIS	Industry	UoB	See CS 1 (below)	Short Task (3-6 months) providing support in the data management process for traceability, semantic annotation & dataset integration.	Started
2	European Material Characterization Council (EMCC)	Supra-national organisation	EwC MU	See CS 2 (below)	Collaboration and Alignment on Ontology Development	Contact Initiated
3	Malta Initiative - Swiss Government, TEMAS	Supra-national organisation	EwC MU	See CS 3 (below)	Discussions initiated on how EwC and partners can support Swiss development under the Malta Initiative (Nano extensions to the OECD guidance for skin sensitization)	Contact Initiated, next meeting scheduled in Switzerland for Q2 2019
4	Blue Frog Scientific	Industry	UoB	See CS 4 (below)	Analysis of services NanoCommons can provide to regulatory consultancies – e.g. benchmarked models, regulatory ontology terms, Klimisch scores or "ToxRTool[1]" scores for datasets etc.	Initial discussions, with follow-up planned in May 2019
5	OECD WPMN	Supra-national organisation	UoB	See CS 5 (below)	Short Task (3-6 months) to annotate existing test guidelines and develop electronic notebook templates, linked to the SOPs for use by researchers and OECD projects.	Pilot project underway for characterisation and ecotoxicity testing
6	EUON / ECHA	Supra-national organisation	UoB	See CS 6 (below)	Integration of NanoCommons into the EUON platform; Analysis of how to utilise the data and tools to identify distinct nanoforms of a NM, and to support grouping and read-across.	Contact initiated & detailed discussions planned for April 2019
7	TEMAS	Consultancy	EwC	See CS 7 (below)	Safety Assessment	Contact initiated
8	CSEM (Swiss Center for Electronics and Microtechnology)	Industry	EwC	See CS 8 (below)	Data Management	Contact initiated – next meeting with VP in April 2019.
9	Contactnanopoint / EMPA	Consultancy	EwC	See CS 9 (below)	Solutions aimed at industry-led needs	Contact initiated – Next F2F meeting on April 29th

Case Study 1: They need help with Data Management (Sensitive data from an industrial point of view). Provide annotations schema. Hierarchical structure of how traceability can be annotated. Create the framework for semantic annotation of traceability data. Implement that in NanoCommons.

Case Study 2: NanoCommons could help them contributing in terms of the Nano part of the database including Tools for characterization.

Case Study 3: Malta initiative extension for skin sensitization.

Case Study 4: Blue Frog Scientific helps other companies with regulatory work. They are interested in Hands-on data and in Regulatory vocabulary. Bilateral consultancy could be a good strategy for sustainability.

Case Study 5: Define vocabulary for OECD guidelines (e.g., characterisation and extension for Nano) that can be used in NanoCommons applications supporting assessment and testing of nanoforms.

Case Study 6: Organization of data. The aim is to solve the problem of representation of assay data involving extensions for nanoforms.

Case Study 7: Partner with risk assessment companies so it would be easier for SMEs to have access to their services.

Case Study 8: Discuss nanotechnology protocol and data management needs of CSEM to determine if we have a problem-solution match to serve. Could be linked to tools for method / assay validation, and/or a support tool to validate methods for regulatory compliance and/or to benchmark data generated via a specific method with data from other related methods.

Case Study 9: NanoCommons and contactnanopoint to work together in defining and solving the needs of industry regarding “regulatory demands” of nanomaterial information provision (e.g., physical chemical property reporting).

5.3 Establishing NanoCommons as a platform for *in silico* nanosafety

NanoCommons came into life as a result of the clear need identified by the nanosafety community represented in large parts but not exclusively by the NanoSafety Cluster to foster Open Data, Open Source and Open Science. Discussions within the NanoSafety Cluster facilitated by the EC, and especially within the context and time scale of the eNanoMapper project (2014-2017) coordinated by DC (now renamed EwC), led to the formation of a sustainability task force led by Barry Hardy (DC and now EwC) with the goal of establishing measures to sustain the work of eNanoMapper. The creation of NanoCommons was one direct outcome of this work and this report provides a milestone to the successful completion of the goal of that task force and a passing of the sustainability work in practice to the NanoCommons partners and associated community.

Data, and tools to process and analyse it, as well as data-driven models, are produced and developed by different publicly-funded projects, but also industry independently; these resources are however rarely shared broadly or successfully and/or are often not based on common database

concepts or harmonised infrastructures limiting the interoperability. To improve on this situation, industry, regulators and policy makers are pushing towards integrated infrastructures and data sharing providing a window of opportunity for initiatives like NanoCommons. Even larger companies are keen to share data for product safety to obtain big data collections suitable for safe-by-design applications and as a side effect of campaigns to improve/restore the trust of the general public in nanomaterials and nano-enabled products, with still not completely explored and exploited application fields. These demands materialize in key political and economic factors, including stricter regulations on materials produced in smaller tonnages and nano-specific regulations and testing guidelines, the drive to reduce animal testing, and the drive towards zero pollution/toxicity policies, which highlight that the nanomaterial toxicology and risk assessment community is in high demand and thus has a high priority for long-term support through a pan-European infrastructure. NanoCommons targets exactly this market and the efforts to sustain its achievements and approaches concentrate on:

1. Identification of possible scenarios for sustainable funding opportunities for guaranteeing the continuation of NanoCommons as a pan-European infrastructure after the end of the project. This will include collaborative efforts with nanomaterial research projects to develop NanoCommons into an Advanced Community having the necessary political and economic impact in a pan-European and international setting and, in this way, to further highlight the importance of nanosafety and its special requirements, which need to be supported by a sustainable infrastructure listed in the ESFRI roadmap.
2. In parallel, we are establishing and intensifying commercialization activities by matching particular resources and services with user and customer interests in the form of needs-based work activities described in the section related to the Case Studies. We also are initiating early discussions with possible providers of additional services, on one hand, and industry groups as long-term users, on the other hand, to ascertain their interest and better understand their requirements for offering and purchasing data and nanosafety/safe-by-design services, respectively. The goal is to probe the commercial potential of the various services and tools via different business models.

NanoCommons follows multiple paths to establish itself as a major European nanosafety service provider, including proactively working to open up opportunities for acquiring additional public funding and in parallel developing a business model for the various services of value to industry, which will together guarantee the long-term sustainability of the infrastructure and its extension into related neighbouring areas like nano safety by design combining material modelling with nanosafety, but also chemical and drug safety and personalised medicine.

The main requirement for being able to sustain the infrastructure beyond the funded project lifetime is to get known across the community as a major European service provider whose services address key market gaps effectively. Only in this way can enough customers - ranging from individual end users to large project consortia requiring data management and knowledge infrastructures but also service providers adding more value to the platform - be attracted. Measures to achieve these goals are presented below in two main areas: 1. Establishing nano toxicology and risk assessment as a major pan-European community with specific infrastructure needs (that can be addressed by

NanoCommons); and 2. Establishing NanoCommons as a place to offer and purchase data, predictive toxicology and risk assessment services.

5.3.1 Establishing nano toxicology and risk assessment as a major pan-European community with specific infrastructure needs

Even if NanoCommons is providing valuable solutions for data harmonization, management and sharing, is giving home to more and more data from different projects and is improving the interoperability between software, further additions and improvements including approaches from all important players in the field will be needed in the future to address all current and upcoming challenges in nanosafety and safe-by-design research and regulations. Therefore, one major task is to constantly enlarge the community by partnering and collaborating with related projects and initiatives, individual researchers and companies including but not limited to the case study partners. Fruitful interactions already exists with the NanoSafety Cluster (as a significant number of NanoCommons partners are members of and actively involved in the NSC (Coordination Team, Steering Group, Dissemination Team, Secretariat members)) and many contributing individual projects, the Nanotechnology Industries Association (NIA), the European Material Characterization Council (EMCC), the European Materials Modelling Council (EMMC), contactpointnano (the Swiss national platform for industrial and regulatory knowledge and expertise on synthetic nanomaterials), the CEINT NanoInformatics Knowledge Commons (NIKC) in the US, the Centre for Sustainable Nanotechnology[4] in the US, as well as OpenRiskNet, which is a more general risk assessment e-infrastructure but with large technology overlap and nano-specific solutions as planned for in its mission and DoA.

However, NanoCommons also has to be fitted into the universe of European infrastructure and establish links to neighbouring communities represented by their own infrastructure projects. Clear candidates, where such links are already well-established, are the chemical, drug and nanomaterial designing and manufacturing communities and the personalised medicine areas including genomics and metabolomics to form a harmonised safe-by-design infrastructure making the tools and services from the individual communities interoperable to each other. However, it is also important to show that risk assessment has specific needs and requirements, e.g. validation of the methods and workflows for their regulatory application including detailed reporting, which have to be supported by specialised tools developed by the community and supported by a well-sustained infrastructure.

Ongoing activities are:

- Becoming an ELIXIR community - work towards this is well underway, with an application lead by NanoCommons and OpenRiskNet partners. Even if the official status was not yet awarded, interactions are already ongoing, e.g. with the ELIXIR TeSS community, and ELIXIR bio.tools and BioSchemas projects and listings of services and training events in the relevant catalogues;

[44] https://susnano.wisc.edu/

- Applying to the next call for ELIXIR industry advisory board members by either a consortium member, a scientific advisory board member or an associated partner to represent the specific requirements of the toxicology community. In the meantime, we will work together with advisory board members from neighbouring disciplines already appointed.

- Engaging with the European Open Science Cloud (EOSC) Strategic Implementation Roadmap activities by registering NanoCommons services in the EOSC hub marketplace and the eInfraCentral catalogue, attending workshops, providing input into consultations and applications to becoming industry board members and/or part of the joint stakeholder forum.

These activities will finally end in an application to become listed on the ESFRI roadmap, which was updated in 2018, and will be updated again in 2020 via consultations over the next year. The next steps to achieve this goal include getting added to the MERIL database and increasing our visibility at research infrastructure conferences and roadmap activities, linking into the relevant national research infrastructure roadmaps and working with the member state delegates to ensure input into and feedback from relevant ESFRI working groups such as the Data, Computing and Digital Research Infrastructures, Innovation, e-Infrastructures Group, and Long-term Sustainability Groups.

5.3.2 Establishing NanoCommons as a place to offer and purchase data, predictive toxicology and risk assessment services

Public funding, as described in section 4.0 and Figure 1 above, can only be one part of the NanoCommons sustainability effort. To guarantee the long-term sustainability including the continuous development, improvement and adaptation to a fast technically and scientifically evolving environment, a business model needs to be established, which supports the goals of Open Data and Open Science but also provides financial resources for the maintenance, improvement and extension of the infrastructure and the provided services.

Options being evaluated currently include:

- **Service Model:** Further development of the core platform is financed by additional commercial services, which the platform provides to the users following the approaches developed in the case studies. Such additional services could include services specifically developed for a client or technical services such as in-house deployment or harmonisation with internal data sources and tools;

- **Data and Software Solution Consortium Membership Model:** The members would be in charge of sustaining the platform and funding its continuous improvement. Starting with the NanoCommons consortium partners, additional companies, institutes but also individuals would become members of the NanoCommons platform paying membership fees based on their legal status and these will be reinvested into new releases;

- **Marketplace platform:** Commercial services will be offered similar to the Service Model. However, not only the platform providers will offer services but third-parties can buy into the platform to increase the feature list. These services would be indexed and offered to the user in exactly the same manner as the services from the platform providers. A share of the generated income for the service provider would go to the platform to cover maintenance, user support and further developments;

- **Virtual Organisation:** Part of the project's exploitation efforts is the establishment of a virtual organisation (VO) business model. In a VO, a number of partners are brought together for a period of time to deliver a product or service and share both resources and costs, which results in lowering of risks and increasing revenue benefits. The VO approach also allows the combination of SME resources into a more extensive competitive entity which can potentially compete on bids for larger projects. Based on previous and ongoing experiences (e.g. ToxHQ: bringing your 21st century safety assessment goals within reach and OpenRiskNet), we are examining the potential application of VOs in sustaining the development of the core infrastructure.

5.4 Next steps in the Development of the Sustainability and Business Plan (during the Project)

This sustainability development plan will be updated annually. During Year 2 of the project (2019) an initial business plan will be developed based on early experiences on obtaining feedback from the market as indicated by the industrial case studies activities described in the section related to the Case Studies.

During Year 1 of the project (2018) we developed and initiated a process for the project for identifying and developing contacts for industrial case studies. This included the innovation of introducing a practice to pay attention to the new regulations of the GDPR (General Data Protection Regulation) when handling contact information.

Factors that we will consider in elaboration of the overall sustainability plan include:

- Criteria for acceptance/integration of data from the different Data Sources (e.g., Nano Safety Cluster projects): how is NanoCommons going to validate if the data is reliable and if its quality is good enough to be included?
- Business plan will include budget developments with an estimation of costs for new product development and delivery;
- Key Performance Indicators (KPI) will be set to:
    1. Determine goals for start dates of a service or tool and how well we execute;
    2. Usage of service or tool;
    3. Funding and income on services selected for development and/or deployment.

- Description of interaction with pilot projects as a set of target users will be included in the sustainability plan – See Section 3.1 - Supra-national organisations addressing nanosafety on stakeholders which includes linkages to the Pilot Lines Network;
- Description of interaction, and a cost model for future projects to access the services provided by NanoCommons will be produced as part of the sustainability plan;
- Specification of a set of NanoCommons value-adding activities (used by different stakeholders) as demonstrated by the Sustainability case studies and demonstration of the value-add arising to the users from these case studies including a cost-benefit analysis of what the services would cost in the future competitive service-based scenario;
- Development of a sustainable business plan for each of the main NanoCommons resources as they are linked into the platform, in order to develop an overall integrated Sustainability plan for the complete platform;
- Development of a Collaboration and/or Virtual Organization business model for NanoCommons;
- Consideration of supporting services such as training, educational and informational services, certification of data quality or certification of data curators, in the business model;
- Including a pricing model in the business plan including consideration of value provided, context of use and feasibility.

5.5 Potential Services to be offered by NanoCommons long term

Identifying and engaging users is very important to achieve the desired sustainability of the NanoCommons e-infrastructure platform: we ask, and will try to answer based on responses from our user test cases, and from early-adopters, the question: which services and tools should be offered to make the platform attractive for a bright and engaged community of users?

The results of (different) surveys will be used for this purpose as established and used in the initial community needs analysis:

https://www.surveymonkey.co.uk/r/PK2KXWW

As a result from the initial survey, NanoCommons has already identified the following services/tools needs of NanoCommons nanosafety users:

- **Design Services** offered by NanoCommons:
  a) Online lab-books, data acquisition
  b) Data curation templates

  Note that Case Study CS5 in Table 2 relates to this type of service, designing the templates and workflows for the OECD SOPs and Test Guidelines.

- **Data Processing and Analysis Services** offered by NanoCommons:
  a) Data cleansing, mining and analysis
  b) Modelling (statistical, mechanistic etc.)

Linking data services with cloud computing capacity to offer on-demand data analysis platforms will present users with a comprehensive environment supporting the full lifecycle of science workflows. Cooperation with the NanoSolveIT H2020 project, which is developing cloud-based solutions, will enable integration of these approaches for NanoCommons users.

- **Data Visualisation and Predictive Toxicity Services** offered by NanoCommons:
  a) Risk assessment tools
  b) Modelling tools

  Note that Case Studies 4, 6 and 7 in Table 2 relate to this type of service, developing predictive models or tools to support risk assessment, dossier preparation, and nanosafety evaluation.
- **Data Storage Services** offered by NanoCommons:
  a) Tool(s) integration
  b) Online data repository and accessibility
  c) Data storage (hardware)

- **Data Management Tools implementing and adhering to FAIR principles.**
  Case Studies 1 and 2 in Table 2 relate to tools / services of this nature, focussing on data management, ontologies and longer-term data storage needs.

These are being addressed by the services currently offered for Transnational Access (TA) or under development for roll-out as TA services in the near future. Readers are referred to the NanoCommons TA portal for more details of these services and tools. For more information/details about the services offered by NanoCommons, and our partner experts, as well as a guideline for the application process, please consult the project website: https://www.nanocommons.eu/ta-access/

5.5.1 Horizon Scanning and potential new tools / services to be developed

Taking a longer term or horizon scanning view, the second phase of development of the sustainability plan will exposure the following potential services, linked to the user case studies identified in Table 2.

- **Nano regulation guide and information services**

Potential initial services envisaged can for example offer solutions to the new regulatory situation providing challenges for companies e.g., satisfying the new regulatory guidance from ECHA with a 2020 deadline. Can NanoCommons provide a regulation guide and offer services to the companies?:

  - Which regulations are actual?
  - What to do/to use to respond to the regulations?
  - How to use the service provided by NanoCommons?
  - Who can help to organize this information?

TEMAS and DC are engaged in discussions to advance this goal (see Case Study 3 in Table 2 above), and UoB are in active discussions also with EUON and ECHA as to development of tools for identification and distinction of nanoforms (Case Study 6).

- **Data Services**

The report "e-Infrastructure for the 21st century – one year later" highlighted Open Access to data (e-Infra 21st). Research advances only through sharing of results, and the value of an investment in research is only maximized through wide use of its results. The European Bioinformatics Institute has estimated that the annual value to users and their funders by making scientific information freely available to the global life sciences community is €1.3 billion per year, equivalent to more than 20 times the direct operational cost of the Institute (Open Research). The e-infrastructure report highlighted that the Horizon 2020 open access data pilot, which encouraged projects to consider open access to data, will also lead to the realisation that it represents an additional cost which has not yet been taken into account for researchers. However, it also presents opportunities for delivery of services to support open data, which NanoCommons and others can leverage.

Among the potential services is long term data preservation, and provision of tools and supports to facilitate this, since data is increasingly recognised as the truly valuable asset of a data-driven economy and hence must be preserved, yet the process of data preservation is not clearly understood by the majority of research communities including nanosafety. This also links to the training services, outlined as possibilities for further development below.

However, the quality of the models and tools that can be developed are highly dependent on the quality and completeness of the underpinning datasets, and thus careful and automated where possible Quality Control and Quality Assurance (QA/QC) steps are needed. Deliverable reports D3.3 - Checklist for use in WP8 / WP9 to support integration of Users data into KB and D4.5 - Workflow and checklist of key information needed from database/dataset owner in order to facilitate integration into KB provide the necessary steps to ensure import of high quality data, and various options for scoring the data completeness and FAIRness are being evaluated and integrated.

In terms of other stakeholder-relevant or stakeholder-specific datasets, a set of criteria will be developed as part of the sustainability plan to consider the question of which data from which sources should feed into the NanoCommons-database. Examples of relevant data sources include:

- OECD WPMN (Malta Initiative and nanoAOP project)
- EFSA: Nano for Intelligent food, food packaging
- 1.1.2020 → New REACH documents including Nano expected
- Nano Guidance for Cosmetics
- Nano Safety Cluster projects
- eNanoMapper database work and its continuation.

- **Training and certification services**

Apart from services, NanoCommons must **teach** its users in order that everybody can fulfil the NanoCommons Principles of FAIRness and Quality Assurance of the generated data through:

a) Training in the generation of FAIR data
   - FAIR = **F**indable, **A**ccessible, **I**nteroperable, **R**e-Usable
   - Open is different from FAIR

b) Data quality assessment:
   - Why it is important and how to do it?
     - Data management plan
     - Tools for doing one
     - Quality assurance plan

c) Certification of users as data curators whose data management skills and datasets meet and exceed the data quality and data completeness criteria established by NanoCommons. This can include certification of datasets themselves as having been integrated into the database by certified curators / data managers.

- **Support Service for users / developers / other e-infrastructures**

To make it easier to use the NanoCommons Platform a **Support Service** will be implemented. Support Services currently in place include the *helpdesk* which provides dedicated persons and a ticketing system to ensure that all queries get addressed or passed to the appropriate experts to be dealt with.

Building on the approaches developed in OpenRiskNet e-infrastructure (https://openrisknet.org/e-infrastructure/developers/) coordinated by partner Edelweiss Connect, a set of services for developers or tools, models and datasets to facilitate their integration into NanoCommons will be developed.

The report "e-Infrastructure for the 21st century – one year later" highlighted the need to accelerate the pace of integration of e-Infrastructure services as a means to increase the adoption of e-Infrastructures by user communities. It encourages existing e-Infrastructure service providers to work more closely together to facilitate the creation of innovative higher-level services (e-Infra). Among the tools offered by NanoCommons that could be offered as services to support integration with other e-infrastructures are the semantic mapping of databases, and the use of multiple-APIs to allow integration. Thus, a next step in the development will be to identify if and how these services can be more widely applied to integrate "nano" data into wider chemical databases, for example.

Close cooperation with other projects involved in modelling of nanosafety is also envisaged, including the recently funded nanoinformatics projects NanoSolveIT and NanoInformaTIX, and the risk governance project RiskGone. Ongoing collaboration with OpenRiskNet (and eNanoMapper) will also be leveraged to co-develop strategies to ensure the sustainability of the generated outputs of the different projects; to leverage the synergies and points in common of all 5 projects and to identify the differences so that it is easier to develop/adapt/align the services and tools offered by each project to the corresponding stakeholders. In the same way that NanoCommons has shared its DMP with the other projects, it will also share its Sustainability plan, and facilitate the co-

development of a long term sustainability plan for all nanosafety and nanoinformatics tools and services. Indeed, the recently completed Nanosafety Cluster Task force on sustainability included some key recommendations that can be taken forward within the scope of this activity, as incorporated upon in this report.

6.0 References

Bastow - Bastow R, Leonelli S: Sustainable digital infrastructure. Although databases and other online resources have become a central tool for biological research, their long-term support and maintenance is far from secure. *EMBO Rep.* 2010;11(10):730–734. 10.1038/embor.2010.145).

Chen - Chen IM, Markowitz VM, Palaniappan K, Szeto E, Chu K, Huang J, Ratner A, Pillay M, Hadjithomas M, Huntemann M, Mikhailova N, Ovchinnikova G, Ivanova NN, Kyrpides NC. Supporting community annotation and user collaboration in the integrated microbial genomes (IMG) system, BMC Genomics. 2016 Apr 26; 17():307).

e-Infra 21st - e-Infratructure for the 21st century – one year later. EIROforum IT Working Group 2nd December 2014
https://archive.embl.org/uploads/r/null/7/f/0/7f070c6cd7a612867bda3f3001776b87b76870f93f74d84108eb74b5c4d7d708/DE_2324_C-EIR-PUBS-B-20.pdf

ECHA Nanoforms - ECHA strategy on substances in nanoforms 48th Meeting of the Management Board 14-15 December 2017
https://echa.europa.eu/documents/10162/2792271/mb_57_2017_echa_strategy_nanoforms_en.pdf/f913484f-9a21-02bc-d386-8cb68d0027a4

Ember C, Hanisch R: Sustaining domain repositories for digital data: A white paper. In *Inter-university Consortium for Political and Social Research (ICPSR)* ICUniversity of Michigan,2013. 10.3886/SustainingDomainRepositoriesDigitalData)

ESFRI Roadmap - Roadmap and Strategy Report on Research Infrastructures
http://roadmap2018.esfri.eu/media/1048/rm2018-part1-20.pdf

FAIR - Report commissioned by the Danish Agency for Science and Higher Education titled "Preliminary analysis: introduction of FAIR data in Denmark" (FAIR)
(https://ufm.dk/en/publications/2018/preliminary-analysis-introduction-of-fair-data-in-denmark)

Ferro - Ferro E, Osella M: Eight business model archetypes for psi re-use. In *Open Data on the Web Workshop* Google Campus, Shoreditch, London, 2013.
https://www.w3.org/2013/04/odw/odw13_submission_27.pdf

Gabella - Gabella C, Durinx C and Appel R. Funding knowledgebases: Towards a sustainable funding model for the UniProt use case https://www.ncbi.nlm.nih.gov/pubmed/29333230

Gabella 2 - Gabella, C, Durinx, C, Appel, R, Funding knowledgebases: Towards a sustainable funding model for the UniProt use case [version 1; peer review: 3 approved] F1000 Research https://f1000research.com/articles/6-2051/v1#ref-15

Lamon - L. Lamon, D. Asturiol, A.Vilchez, R. Ruperez-Illescas, J. Cabellos, A. Richarz, A. Worth. Computational models for the assessment of manufactured nanomaterials: Development of model reporting standards and mapping of the model landscape. Computational Toxicology. Volume 9, February 2019, Pages 143-151. https://www.sciencedirect.com/science/article/pii/S2468111318300847?via%3Dihub

Maron - Maron NL: A guide to the best revenue models and funding sources for your digital resources. 2014. https://sr.ithaka.org/wp-content/uploads/2015/08/Jisc_Report_032614.pdf

Open Research - Open Research Data in H2020 https://ec.europa.eu/research/press/2016/pdf/opendata-infographic_072016.pdf

Poux - Poux S, Arighi CN, Magrane M, *et al.* : On expert curation and scalability: Uniprotkb/swiss-prot as a case study. *Bioinformatics.* 2017;33(21):3454–3460. 10.1093/bioinformatics/btx439 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5860168/