RSC Sustainability

PAPER

Cite this: *RSC Sustainability*, 2024, **2**, 3464

The FAIR principles as a key enabler to operationalize safe and sustainable by design approaches

Achilleas Karakoltzidis,[ab] Chiara Laura Battistelli,[c] Cecilia Bossa,[c] Evert A. Bouman,[d] Irantzu Garmendia Aguirre,[e] Ivo Iavicoli,[f] Maryam Zare Jeddi,[g] Spyros Karakitsios,[ab] Veruscka Leso,[f] Magnus Løfstedt,[h] Barbara Magagna,[i] Denis Sarigiannis,[ablm] Erik Schultes,[ijk] Lya G. Soeteman-Hernández,[g] Vrishali Subramanian[g] and Penny Nymark[*n]

Safe and sustainable development of chemicals, (advanced) materials, and products is at the heart of achieving a healthy future environment in line with the European Green Deal and the Chemicals Strategy for Sustainability. Recently, the Joint Research Center (JRC) of the European Commission (EC) developed the safe and sustainable by design (SSbD) framework for definition of criteria and evaluation procedure proposed to be established in Research and Innovation (R&I) activities. The framework aims to support the design of chemicals, materials and products that provide desirable functions (or services), while simultaneously minimizing the risk for harmful impacts to human health and the environment. While many industrial sectors already consider such aspects during R&I, the framework aims to harmonize safety and sustainability assessment across diverse sectors and innovation strategies to meet the mentioned overarching policy goals. A cornerstone to successfully implement and operationalize the SSbD framework lies in the availability of high-quality data and tools, and their interoperability, aspects which also play a key role in ensuring transparency and thereby trust in the assessment outcomes. Availability of data and tools depend on their machine-actionability in terms of findability, accessibility, interoperability, and reusability, in line with the FAIR principles. The principles were developed in order to harmonize digitalization across all data domains, supporting unanticipated data-driven "seamless" integration of information and generation of new knowledge. Here we discuss the essentiality of FAIR data and tools to operationalize SSbD providing views and examples of activities within the European Partnership for the Assessment of Risks from Chemicals (PARC). The discussion covers five areas previously brought up in relation to the SSbD framework, and which are highly dependent on implementation of the FAIR principles; (i) digitalization to leverage innovation towards a green transition; (ii) existing data sources and their interoperability; (iii) navigating SSbD with data from new scientific developments (iv) transparency and trust through automated assessment of data quality and uncertainty; and (v) "seamless" integration of SSbD tools.

Received 10th April 2024
Accepted 15th September 2024

DOI: 10.1039/d4su00171k

rsc.li/rscsus

Sustainability spotlight

In an era characterized by escalating environmental challenges and mounting concerns regarding public health, it becomes imperative to embed safety and sustainability principles across all stages of innovation to ensure environmentally friendly chemicals, materials and products. This study points to the importance of leveraging the FAIR principles to support efficient machine-actionable data and tool reuse coupled to the recent European Commission-recommended framework for Safe and Sustainable by Design (SSbD) approaches. Only through harmonized and digitalized data-driven assessment of

[a]HERACLES Research Center on the Exposome and Health, Center for Interdisciplinary Research and Innovation, Aristotle University of Thessaloniki, Thessaloniki, Greece
[b]Environmental Engineering Laboratory, Department of Chemical Engineering, Aristotle University of Thessaloniki, Thessaloniki, Greece
[c]Environment and Health Department, Istituto Superiore di Sanità, Rome, Italy
[d]Environmental Impacts and Sustainability, NILU, Kjeller, Norway
[e]European Commission, Joint Research Centre (JRC), Ispra, Italy
[f]University of Naples Federico II, Naples, Italy
[g]National Institute for Public Health and the Environment (RIVM), Bilthoven, Netherlands
[h]European Environment Agency, Copenhagen, Denmark
[i]GO FAIR Foundation, Leiden, Netherlands
[j]Institute for FAIR and Equitable Science, Leiden, The Netherlands
[k]Academic Centre for Drug Research, Leiden University, The Netherlands
[l]National Hellenic Research Foundation, Athens, Greece
[m]University School of Advanced Study IUSS, Pavia, Italy
[n]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. E-mail: penny.nymark@ki.se

human health and environmental impacts of emerging technologies can we foster sustainable industrial innovation (SDG 9) and responsible production (SDG 12), in order to ensure the safety and well-being of end-users (SDG 3) and the environment (SDGs 6, 14 and 15).

Background and context

Digitalization refers to the process of converting data (as in raw data points) and information (as in interpreted data or knowledge) into a digital format that is readable by a computer.[1] The FAIR principles were coined in 2016 to guide how to integrate or to harmonize digitalization of data and software (in the following referred to as tools) so that they become findable, accessible, interoperable, and reusable by both machines and humans.[2,3] The current wave of data and information generation both within science as well as in industrial and regulatory environments requires machines to optimize and reach the full potential of the information generated. With the recent rise of novel artificial intelligence (AI) tools, the world has seen a new level of support that can be provided to human activities. However, machine-driven tools are only as good as the data they can access and reuse. In addition, humans can only trust the outcomes of machines if the process is transparent and clear about the associated uncertainties, often referred to as explainable AI.[4]

The safe and sustainable by design (SSbD) framework proposed by the European Joint Research Centre (JRC) was recommended by the European Commission as a strong piece of the puzzle to reach the European environmental policy goals set out in the Green Deal and the Chemicals Strategy for Sustainability (CSS).[5–8] Worth mentioning is also the endeavour for circular economy in support of the CSS and SSbD challenges.[9] The JRC framework was recently found to be the most comprehensive description of SSbD to date and serves as a basis for the discussions in the current paper.[10] The framework can be referred to as a pre-market approach taken during research and innovation (R&I) to support harmonized design, development, production, and use of chemicals, materials, and products focusing on providing desirable functions (or services), while simultaneously minimizing harmful impacts to human health and the environment, in particular groups of chemicals likely to be (eco)toxic, persistent, bio-accumulative or mobile.[11] The approach describes five steps addressing: (1) hazard of chemicals/materials, (2) occupational safety and health, (3) the human and environmental aspects during the final application phase of chemicals/materials, (4) environmental sustainability, and (5) socio-economic sustainability.[11] Data and competencies have recently been found to be among five important building blocks required to implement SSbD in practice.[10] Data enables reliability, traceability and transparency, while competencies

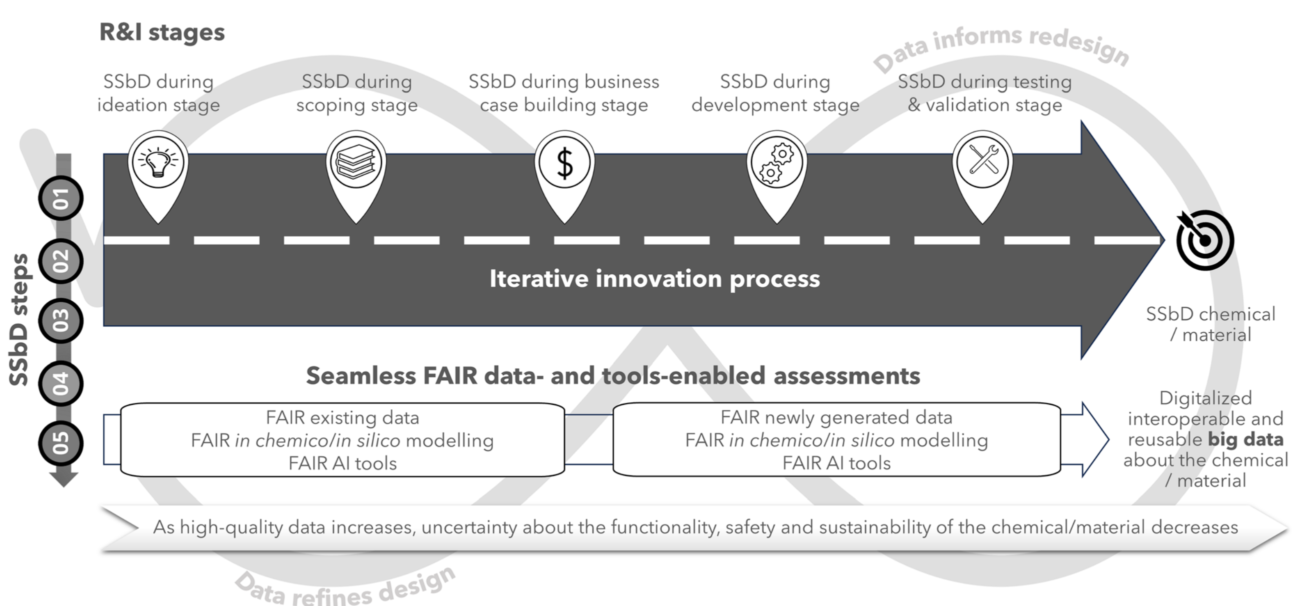

Fig. 1 Overview of the essentiality of FAIR data and tools to operationalize SSbD along the research and innovation (R&I) process. Harmonized digitalization in line with the FAIR principles allows for consideration of all relevant data domains and tools during R&I, including for targeted design of functionality and for assessment of safety and sustainability parameters, supporting data-driven integration of information and ultimately unanticipated discoveries towards a green transition. In the context of the figure, the following terms are defined as FAIR existing data = digitalized and FAIR data gathered from SSbD relevant databases useful for *in chemico*/*in silico* modelling and/or data generated in previous stages of the R&I process; FAIR *in chemico*/*in silico* modelling = computational modelling using FAIR models and tools; FAIR AI tools = highly advanced data-driven FAIR and explainable artificial intelligence tools supporting decision-making.

Table 1 Overview of the original FAIR principles and examples of the needs to ensure their successful implementation into SSbD. The **bold** marked text indicates aspects related to the so-called blue FAIR principles (as presented in Schultes[13]) which require extensive discussions and social agreements on the content-related, domain-relevant aspects of the FAIR principles. These discussions support the implementation of the necessary technicalities of FAIR orchestration within *e.g.* the SSbD framework

FAIR principles	Implementation to SSbD	Social agreements needed
Findability		
(F1) (Meta)data are assigned a globally unique and **persistent** identifier	Globally unique and **persistent** identifiers (*e.g.*, DOI) for datasets and tools support searchability in the SSbD data collection, organization, and integration phases	The need and level of persistence for identifiers used for SSbD-relevant datasets and tools needs to be agreed on
(F2) Data are described with **rich** metadata (defined by r1 below)	Well described datasets that are **semantically** annotated with a plethora of **refined keywords relevant to SSbD** support inclusion and integration [also relates to I2]	Agreements on minimum information requirements for SSbD-relevant metadata are needed
	Comprehensive documentation, including data dictionaries, codebooks, and README files that explain the **structure**, **variables**, and usage, support transparency and the probability of uptake into the SSbD assessment [also relates to F3–R1]	SSbD community-endorsed metadata schemas are needed
(F3) Metadata clearly and explicitly include the identifier of the data they describe	Embedded models including metadata annotations of datasets support finetuning of findable data and enhances the potential of compatibility among datasets for SSbD assessment [also relates to F4–R1.2]	—
(F4) (Meta)data are registered or indexed in a searchable resource	Interconnections with peer reviewed databases through high-performance APIs support increased numbers of potential data sources considered within the SSbD framework [also relates to I1–I3]	—
Accessibility		
(A1) (Meta)data are retrievable by their identifier using a standardised communications protocol	Standard communication protocols and data exchange protocols (*e.g.*, REST, OData) facilitate data exchange and integration and assessment of unanticipated dataset's relevance to SSbD	—
(A1.1) The protocol is open, free, and universally implementable	Open, free, and universally implementable protocols support broad uptake of data and tools within the SSbD framework	—
(A1.2) The protocol allows for an authentication and authorisation procedure, **where necessary**	Login systems to access (meta)data with authentication or authorization methods **to manage user access** allow for efficient machine-driven data integration within the SSbD pipeline [also relates to A2]	Agreements on the need for authentication and authorization and to which extent (to the level of metadata or data) are needed. In addition, agreements on the stability/sustainability of hosting platforms/repositories is needed
	Dataset hosting on **stable** and accessible platforms or repositories, with **well-defined access policies** with the delivery of high-performance APIs support effective SSbD processes [also relates to I1–R1.3]	
(A2) Metadata are **accessible**, even when the data are no longer available	**Accessible**, and preferably **open to the extent possible**, metadata supports assessment of dataset's relevance for SSbD	Agreements on the necessary level(s) of openness and persistence for metadata are needed
Interoperability		
(I1) (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation	Standard data schemas and formats, use of freely accessible data formats (XML, JSON-LD or RDF), allow for broad and equal processing by everyone and anyone within the SSbD community	—
	Integration of data through shared programming libraries and packages for popular programming languages as well as use of broadly applicable APIs increase interoperability between SSbD-relevant fields [also relates to A1]	
(I2) (Meta)data use **vocabularies** that follow FAIR principles	**Semantic artefacts, *e.g.* ontologies and controlled lists of terms, that use identifiers for referencing defined concepts** in relevant (meta)data and codebooks support broad interoperability between SSbD-relevant fields. [Also relates to R1]	Agreements on SSbD-relevant ontologies and vocabularies are needed

Table 1 (*Contd.*)

FAIR principles	Implementation to SSbD	Social agreements needed
(I3) (Meta)data include **qualified references** to other (meta)data	**Linked (meta)data** based on standard metadata schemas and semantic models, enable exploration of interconnections and dependencies with other unanticipated SSbD-relevant data sources [also relates to F2]	Agreements on SSbD-relevant metadata schemas, ontologies and minimum information requirements that support inclusion of qualified references and allow for linkage to unanticipated data sources for SSbD, are needed
Reusability		
(R1) (Meta)data are richly described with a **plurality of accurate and relevant attributes**	**User support** and **documentation** for navigating and accessing datasets, including **tutorials or FAQs**, supports broad and harmonized reusability of data in SSbD activities [also relates to I1–A1]	Agreements on the level of SSbD-relevant documentation and attributes is needed
(R1.1) (Meta)data are released with a clear and accessible data **usage license**	Clear specification of **licensing terms and restrictions** ensures appropriate reuse and sharing of data and results from and within the SSbD assessment [also relates to R1.3]	Agreements on licensing policies across SSbD-relevant platforms and repositories is needed
(R1.2) (Meta)data are associated with **detailed provenance**	Incorporation of blockchain technologies supports inclusion of **rich provenance (meta)data** providing outlines of data owners, and users within the SSbD context [also relates to F2] Recorded lineage of the dataset, including information on **how it was collected, processed, and updated** increases user awareness and **transparency promotion**	Discussions regarding the necessary level of provenance detail are needed
(R1.3) (Meta)data meet **domain-relevant community standards**	Data transformation tools assist users in converting data to an SSbD-relevant **format** compatible with their systems based on the agreed community standards **Encouraged user feedback and engagement** addresses issues over time, continuously improves overall FAIRness (and quality) of data, and thus the reliability of the SSbD assessment	Agreements on SSbD-relevant community standards for (meta)data formats and structure are needed

are supported by easy-to-use accessible tools, tutorials, platforms and training.[10]

Indeed, the lack of data has been noted as a major issue in all value chains where SSbD has currently been considered.[10] Data availability is especially crucial at the early stages of R&I when data on the chemical/material at hand is scarce for obvious reasons. At these stages, access to tools that can interoperate with existing data and information to model and/or predict functionality, safety and sustainability becomes valuable. At later stages newly generated data using cost-efficient screening technologies becomes relevant and overall accumulates increasingly bigger data about the chemical/material at hand (as reviewed recently from the safety perspective by Nymark, *et al.*[12]). The increasingly bigger volumes of data support decreased uncertainties about the functionality, safety, and sustainability of the chemical/material, and in turn support efficient assessment of trade-offs between the SSbD dimensions, which has been identified as crucial in order to avoid trade-offs on specific safety or sustainability aspects due to predefined cut off criteria.[10] However, to function seamlessly in concert, data and tools need to be FAIR. See Fig. 1 for overview of the seamless support that FAIR data and tools can provide for the SSbD approach. In the figure, the SSbD steps (vertically to the left) happen along each stage in the iterative R&I process (horizontally), and each stage is coupled to increasing amounts of FAIR(ified) data, first existing gathered data, and at later stages newly generated data. As high-quality data accumulates and becomes increasingly bigger along the stages of the R&I process, uncertainty about the functionality, safety, and sustainability of a chemical/material decreases by design. Overall, existing, and newly generated data refines design, while the increasingly bigger and comparable data gathered along the SSbD process iteratively informs redesign, as depicted by the infinity arrow.

The FAIR principles were designed to be aspirational and hence, do not provide precise guidance for direct implementation into specific domains. Thus, successful implementation of the FAIR principles into the SSbD domain requires consideration of specific needs within the domain and includes both social and technical aspects.[13] The social aspects of FAIRification[14] involve agreements within and across specific domains regarding *e.g.*, the use of standards, metadata templates, controlled vocabularies (*e.g.* ontologies), and authentication/authorization requirements, while the technical aspects of so-called FAIR orchestration involve broadly applicable general data management solutions allowing for data and tools to become susceptible to reuse in unexpected manners. Currently, the social aspects require dialogue to advance implementation within the SSbD domain. Examples include discussions regarding domain-relevant minimum information requirements, structure of (meta)data schemas, vocabularies and requirements relating to persistence, openness, and licensing.

It is especially worth highlighting that FAIR data principles do not inherently necessitate openness in data access with unrestricted use. On the other hand, metadata can be openly available without jeopardizing data that necessitates restriction promoting data findability as will be discussed in detail later. Thus, FAIR data can still be subject to varying degrees of accessibility, encompassing access controls and licensing agreements, which influences the extent to which it can be utilized or disseminated.[15] Table 1 provides an overview of the original FAIR principles and raises some examples of social aspects requiring discussion within the SSbD domain.

The objective of this paper is to provide further insight into the importance of the FAIR principles for operationalizing SSbD,[11] and why the principles should play a central role in the development of SSbD toolboxes to allow for seamless integration of data and tools. The paper covers five areas previously brought up in relation to the SSbD framework proposed by the JRC,[11] and which are highly dependent on the implementation of the FAIR principles; (i) digitalization to leverage innovation towards an effective data-driven green transition; (ii) existing data sources and the quest for interoperability; (iii) navigating SSbD with data from New Approach Methodologies (NAMs); (iv) transparency and trust through (semi)automated assessment of data quality and uncertainty; and (v) "seamless" integration of SSbD tools. In addition, we provide views and examples of activities within the European Partnership for the Assessment of Risks from Chemicals (PARC), as well as other ongoing EU projects.

Digitalization to leverage innovation towards an effective data-driven green transition

Digitalization is a key component in the green transition acting as a catalyst for increased efficiency, sustainability, and R&I across multiple sectors and supports the combination of data from numerous sources such as complex supply chains, stakeholders, business models, human and environmental health, and sustainability.[1,16,17] By leveraging digital technologies such as IoT (Internet of Things),[18,19] AI,[20] cloud computing,[21] and data analytics,[22,23] individual safety and sustainability components (*e.g.*, hazard, exposure, resource usage, environmental impact, emissions, *etc.*) can be optimized.[24]

Nevertheless, the process of digitalization demands considerable resources, including economic and computational resources, manpower, and time, as well as the implementation of harmonized FAIRification processes. It is propelled by technical solutions supporting FAIRification, as exemplified in Table 1 and illustrated in the following example: Achieving data reusability often necessitates assessors, who may not be the original data providers, to access information through numerous methodological frameworks. Traditionally, data has to be downloaded, pre-processed, and opened in dedicated stand-alone software (tools) before advancing to the subsequent re-use step, as demonstrated in a recent case study conducted by the JRC.[25] The resultant outcomes then have to be saved in a repository. Each of these steps are variably performed by a wide range of data reusers with widely different profiles, expertise and aims. However, converting unstructured data into structured format is fundamental for harmonized organization of information within the FAIRification framework, which is constitutional to the analytical process.[26]

Advancements in digitalization, coupled with the availability of network-based digital infrastructures, now facilitate the integration and operationalization of raw datasets and FAIR tools. This modular, service-based architecture permits FAIR data transfer between applications through mechanisms such as REST APIs (Representational State Transfer Application Programming Interfaces), grounded in concepts such as uniform interfaces and client-server decoupling. When these standard communication protocols are combined with network-based services, workflows become significantly more streamlined, since both inputs and outputs adhere to a machine-readable FAIR standard, and all necessary processing information is included in the standardized communication and data transfer between services.

However, the social agreements on how to harmonize digitalization within the SSbD community are lacking. Numerous digital solutions effectively supporting SSbD have been proposed, including digital twins that enable real-time monitoring, simulation, and optimization prior to the development stage,[27,28] digital product passes,[29] traceable material loops,[30] and the establishment of a European common data platform with the aim of facilitating the sharing, access, and re-use of information on chemicals.[5,31,32] These solutions result in a significant enhancement of economic and social operational efficiency.[33,34] Furthermore, a plethora of companies have embraced "Industry 4.0" principles which include most of the solutions described.[35–37] As a result, such companies have achieved increased productivity and efficiency,[16] especially when incorporating predictive maintenance into industrial processes.[38] Overall, the green transition necessitates integration of data from diverse sources, emphasizing the significant role of digital technologies in shaping a more sustainable future.[17]

Existing data sources and the quest for interoperability

Today, available information relevant to SSbD is often "unFAIR". For example, toxicological data is often stored in (confidential) text documents and (closed) servers, such as data generated within EU projects, and thus not accessible to external users.[39–42] During the past decade, attempts have been made to store legacy and newly generated data in more FAIR databases. However, while on the right path, these endeavors have often resulted in lost raw data and information, due to aggregation in *ad hoc* repositories by others than the original data generators.[42,43] Thus, the original nature of these types of data is often lost in translation transition. These problems become even more compelling when the criteria for results interpretation change (*e.g.*, the acquisition of new knowledge) and poorly reported data become unusable. A substantial amount of information from many different types of sources is relevant when performing an SSbD assessment making data retrieval ineffective and time consuming. The complexity of these endeavours is exemplified in several ongoing nanomaterial and pharmacological projects focused on SSbD (to mention a few: SSbD4CHEM†, CHIASMA‡, CheMatSustain§, PINK¶, SUNSHINE‖, PREMIER**, TRANSPHARM††) have adopted the FAIR principles and aim to implement FAIR-ification approaches in parallel with the development and generation of both tools and data.[39,44]

As the awareness of the FAIR principles, as a requirement for good scientific practice, has grown in recent years,[45] FAIR data requirements are now requested in (EU) research projects and early adoption of the FAIR principles is strongly encouraged. Nevertheless, since the FAIR principles are aspirational in nature and do not provide stringent guidance on how to make specific types of data FAIR, their practical implementation can vary greatly, resulting in datasets with highly different levels of FAIRness, especially across different domains.[46,47] Diverse tools have been developed to assess the compliance of the datasets with the FAIR principles and FAIR implementation networks have been established to support broad discussions aimed at harmonization within specific communities and domains.[44,48] So-called FAIR maturity indicators provide an objective quantification of the FAIRness level achieved, and practical guidance for its improvement.[49] These tools help differentiate the needs of two distinct areas of expertise addressed: the data science and the data content, which correspond to the technical and social aspects of the FAIR principles, respectively (*cf.* Table 1).[13,48] Given that these two areas require different and often very specialized expertise, and that FAIR implementation involves distinct but highly interconnected activities, efficient communication between the two is needed.[50] Whilst data science implies Information Technology (IT) skills, and is often agnostic to the actual data content, data domain expertise is needed to identify and implement specialized domain requirements, *i.e.* building on social agreements. Applying data and tools based on standards along with appropriate domain-relevant content standards and accessible rich metadata that uses harmonised terminology supports interoperability thereby avoiding the need for manual transformation and/or mapping, and reduce the time needed for the SSbD assessments.

A noteworthy example of how these characteristics can be implemented in a single harmonized approach, is the QSAR Toolbox‡‡ which is a software application designed to support *in silico*-based hazard assessment of chemicals, incorporating interoperable data and tools from numerous sources. Another example is the bioinformatics community, which has long developed a broad suite of interoperable data and tools in the form of repositories with omics data and R-script tools.[51] Similar characteristics are important for SSbD-relevant data and tools, which should be capable of raising red flags based on existing data in terms of any of the SSbD dimensions (functionality, safety, sustainability) at early stages of R&I, and preferably simultaneously direct decision-making towards more promising alternatives, allowing for iteration during the R&I process (*cf.* Fig. 1).

Navigating SSbD with data from new approach methodologies (NAMs): an iterative interplay

One of the needs for operationalizing SSbD is the availability of suitable and efficient methods for testing and assessing the different parameters/information requirements associated with each SSbD dimension.[25] The conventional methodologies for chemical risk assessment are evolving towards frameworks consisting of cost-effective New Approach Methodologies (NAMs), which serve as a good example for describing the needs for methods to operationalize SSbD. NAMs aim to replace, reduce and refine the need for lengthy animal testing to meet regulatory requirements.[39] NAMs encompass a range of technologies, methods, and strategies that can provide information on a wide variety of parameters relevant to environmental and human health risk assessment,[8] including approaches for grouping and read across, exposure assessment and modelling, *in silico* modelling of physicochemical structure and hazard data, *in vitro* high-throughput and high-content screening assays, dose–response assessments and modelling, analyses of biological processes and toxicity pathways, kinetics and dose extrapolation, and consideration of relevant exposure levels and biomarker endpoints, including also AI.[4,12,26,52,53] Having a robust collection of existing knowledge for validation is important for the development of NAMs aimed at predicting toxicity pathways and outcomes. In principle, AI modeling has the potential to enhance the evidence foundation of such NAMs.[26]

The novelty of NAMs is related to their novel application to regulatory decision-making,[54] but NAMs are also considered to provide significant support to reduce uncertainties regarding the safety (and in some cases sustainability) parameters during R&I processes.[12] NAMs can be used alone, or in combination in Integrated Approaches to Testing and Assessment (IATA) or Defined Approaches (DA) providing sufficient information with higher confidence to evaluate the risk for adverse effect on human health and/or the environment.[55] A first successful example is the DA for skin sensitization, recently adopted as an OECD guideline, demonstrating that the limitation of a single *in vitro* method can be overcome by using several NAMs in a specific combination, and the resulting data are interpreted using a fixed data interpretation procedure.[56]

† https://www.nanosafetycluster.eu/nsc-overview/nsc-structure/ongoing-projects/ssbd4chem/ (accessed December 2023).

‡ https://chiasma-project.eu/ (accessed December 2023).

§ https://www.nanosafetycluster.eu/nsc-overview/nsc-structure/ongoing-projects/chematsustain/ (accessed December 2023).

¶ https://pink-project.eu/ (accessed December 2023).

‖ https://www.h2020sunshine.eu/ (accessed December 2023).

** https://imi-premier.eu/ (accessed December 2023).

†† https://transforming-pharma.eu/ (accessed December 2023).

‡‡ https://qsartoolbox.org/ (accessed December 2023).

The use of NAMs has been prioritized due to their ability to improve quality measures in data generation such as relevance, sensitivity, accuracy, depth of understanding and harmonized reporting, which together brings useful reference data of high quality.[12] For instance, NAMs were applied for risk assessment of the substance tebufenpyrad in the work of Alimohammadi, *et al.*[57] showing their importance in informing regulatory decisions to safeguard human health. Inclusion of NAM-derived data and screening for possible hazardous properties at an early stage of the R&I process enables the assessment of chemicals currently not covered by REACH or other regulations, such as advanced materials.[11,58] For new chemicals and materials, NAMs are essential for screening since *in vivo* tests are too costly and time consuming to be considered early in the R&I process. In addition, the support and promotion of the use of NAMs in the SSbD approach is especially relevant, as the potential benefits gained from building up high-quality interoperable big data resources about all chemicals/materials, processes and products developed, that iteratively informs and improves redesign (*cf.* Fig. 1), is significant.[12]

However, given the substantial volume, variety, and velocity of data generated by NAMs, it is imperative to ensure that these types of data and tools are made FAIR, and for their usefulness to specifically SSbD, the FAIRification must be aligned with the associated social agreements within the domain. Therefore, standardized templates for NAMs data can promote harmonization of the collection, storing, and sharing of information among end-users, and contribute to data consistency and quality.[59] To this end, the OECD has promoted the OECD Harmonized Templates (OHT) tailored for documenting information relevant to the intrinsic properties of chemicals, encompassing effects on both human health and the environment§§. Notably, a new OHT has been recently implemented, particularly relevant for reporting of data from NAMs, namely OHT 201 on intermediate effects[59] with the aim to harmonize the collection of mechanistic information. However, FAIRification of NAMs data remains a crucial goal to promote confidence and advance their reusability especially for SSbD purposes. Successful operationalization of the SSbD framework relies on transparent assessment processes supported by FAIR data generated using NAMs, along the entire life cycle of the chemical or material. For example, it is worth mentioning the Adverse Outcome Pathway (AOP) framework, which, if FAIR itself, has the potential to serve as a transparent platform for improved visibility and increased trust in NAMs data.[31] Notably, the previously mentioned OECD DA for skin sensitization builds on an AOP, however, the handling of the data and results is poorly described in the guideline and would benefit from further development towards including guidance on how to improve data FAIRification.

Finally, it is worth mentioning that several NAMs currently applied for safety assessments align with *in silico* modelling and high-throughput screening approaches used in material design approaches to assess and predict functionality.[12,60] Thus, it becomes interesting to speculate that also the sustainability dimension could be addressed through similar approaches. Overall, the SSbD toolbox will need to be flexible towards interoperability with NAMs data and tools, including future new types of NAM data, considering the extended view of NAMs also covering methods for addressing both the functionality and sustainability dimensions.

Transparency and trust through (semi) automated data quality and uncertainty assessment

The SSbD approach is dependent on transparency and trust in the assessment outcomes. Transparency and trust depend on openness about the interrelated concepts of data quality and uncertainties, and dealing with these is an inherent part of chemical safety- and sustainability assessments. For example, in terms of the quality of risk assessment data, different approaches can be used depending on the type of data, including *e.g.* the Klimisch scoring system (*in vitro* and *in vivo* studies) and OECD QSAR assessment framework (*in silico* predictions).[61,62] The evaluation of uncertainties then requires access to sufficient information on the individual study, model, or prediction to perform an independent evaluation. Depending on the information available (or lack thereof) different scorings are obtained for reliability (including both methodological and reporting quality) and relevance (to the problem at hand) of the data.[63,64] These types of scorings are most trustworthy if they are performed in harmonized fashion and include transparency regarding how the scores were obtained. Novel flexible and transparent approaches for data quality scoring have recently been suggested for non-validated NAMs data in order to support its application in rigorous regulatory risk assessments, and also provide opportunities within SSbD approaches.[63,65]

However, an all-encompassing SSbD assessment requires the integration of large amounts of data points of different levels of quality from multiple sources. FAIR data (and tools) support this process not only by expanding the amount of relevant data that can be (automatically) retrieved, but also to increase transparency regarding data quality and associated uncertainties which is crucial to achieve trustworthy results. The large amounts of data that are relevant to an SSbD assessment creates broad domain-specific challenges when it comes to assessment of data quality and uncertainties. Worth mentioning is the challenge that different levels of uncertainties may be tolerated at the beginning of the SSbD process (acceptance for higher uncertainties) as compared to later stages (requirements of very low uncertainties) (*cf.* Fig. 1). Thus, quality assessment can be context-specific depending *e.g.* on the problem formulation at hand. For this reason, it is important to note that FAIR data not only serve to increase the integration of data into SSbD assessments, but also provide basis for further developments of (semi)automated data quality and uncertainty assessments (based on the machine-actionability that FAIRification provides).[39] Such developments will be of paramount importance *e.g.* for trade-off assessments between the different SSbD

§§ https://www.oecd.org/ehs/templates/ (accessed December 2023).

dimensions (*i.e.* functionality, safety, sustainability *etc.*), and for comparative assessments between mature/on the market data-rich chemicals and new/under development data-poor alternatives.[10]

The lack of FAIR data, which also hinders their quality assessment, reverberates on the *in silico* methodologies that could be used in case of lack of experimental evidence in the R&I stages of the SSbD framework, *e.g.* grouping and read across approaches. The application of such modeling approaches relies on existing good quality data and suffers from the same limitations as the data used for its implementation. The evaluation of the models and their predictions remains a crucial point for their exploitation in general and specifically in the SSbD framework.[66] A big step forward in the direction of the harmonization of model evaluation and of the improvement of their regulatory acceptability is represented by the recent release of the OECD QSAR Assessment Framework.[62] In parallel, to facilitate and harmonize models sharing and exchange, adaptation of FAIR criteria specifically for models has been proposed.[3,67,68]

"Seamless" integration of data and tools during SSbD: is it feasible?

The term "seamless" has frequently been used to envision future computational toolboxes and platforms where data and tools work together to provide user-friendly solutions to a variety of challenges within regulatory and industrial chemical and material risk assessment approaches.[4,69–72] Indeed, the seamless integration of tools and data stands as the next significant milestone to operationalize SSbD. From a computational standpoint, SSbD represents a collection of tools and data, intricately combined to formulate, and execute an approach aligning with the EU's chemicals policy and its established objectives through the year 2050.[5–7] To illustrate the procedure, the common practise currently is that data extracted by tool A is, for instance, translated for tool B, and the analysis continues through tools C–D, and so forth, until reaching a final SSbD score, which supports a decision at a certain point during the R&I process. This partially manual process, from conceptualization to final analysis, has proved inconvenient, resource intensive and is error prone.[2] The complexity is heightened by the multitude of tools available in the scientific community, each designed to address specific SSbD steps, requiring significant expertise and experience.

These challenges can be significantly supported through the adoption and application of the FAIR principles allowing for trustworthy, explainable (*i.e.* broadly understandable), data-driven, and machine-generated results supporting comparable, transparent, and justifiable decisions.[4] The FAIR principles allow the inputs and outputs of diverse tools to seamlessly connect (due to their machine-actionability) and exchange data through automated systems, all within a fully transparent framework, resulting in the integration of highly informative and intuitive visualization of all available data and tools being employed (*e.g.* the importance of visualization in data management and presentation through AOPs is discussed in Wittwehr, *et al.*[31]). When all tools calculate results within a consistent range, there is a substantial increase in the overall reliability, and precision of the adopted methodology, *i.e.*, leading to higher quality results. Simultaneously, the number of tool combinations capable of producing the desired outcome increases. This expansion includes determining the specific tools to be used, highlighting the establishment of the methodology as a crucial factor for achieving SSbD chemicals, materials, and products and ensuring reliable and harmonized decision-making. It can be expected that as more tools are incorporated to derive a result, and these results converge or vary within a common range, more precise estimates of the degree of uncertainty arising from the calculations will be possible, supporting decision-making significantly.[73]

Overall, it is worth noting that the seamless integration of tools does not focus only on the tools themselves but also enhances the management of the available and newly generated data. Within such a framework, data can be directly absorbed and categorized within the tools, simplifying the analysis process, and facilitating its application in the field of use. Indeed, this step is evidently a starting point to establishing a robust framework for FAIR data-driven decision-making within the context of SSbD. Ultimately, such an endeavour substantially diminishes the complexity of calculations and enhances the user experience. This is particularly crucial and important for the industry, empowering stakeholders to take ownership of the process and become integral participants in the entire undertaking throughout value chains.[10]

The alignment of the FAIR principles with the SSbD process is illustrated in Fig. 2, demonstrating that a lack of FAIR data and tools hinders R&I and SSbD. Along the SSbD steps, existing FAIR data and tools, including from NAMs and from diverse sources with sustainability and socioeconomic data, are a prerequisite to allow for a reliable progression to the subsequent steps of the analysis. A lack of FAIR data and tools significantly hinders the progression as demonstrated in the JRC SSbD framework case studies.[25] Overall, the first step is dependent on availability of data relevant to assessment of hazard, through *e.g.* grouping and read across approaches, which is often not available for new chemicals/materials. The second step focuses on assessment of occupational safety and health, where potentially sensitive industrial data becomes relevant. The third step addresses the human and environmental aspects during the final application phase of chemicals/materials, where the General Data Protection Regulation[74] may be relevant to consider. The fourth step focuses on diverse life cycle aspects of the chemical/material at hand and requires extensive understanding of and insight into the chemical's/material's application area. At this point and in the fifth step, where the socioeconomic sustainability of the chemical/material is assessed, the assessment of the final product and its suitability for market distribution is significantly aided by the availability of findable, if possible open, and reusable (meta) data. The final decisions regarding whether to proceed to the next stage with the chemical/material in the R&I process, depend on the robustness and trustworthiness (in terms of

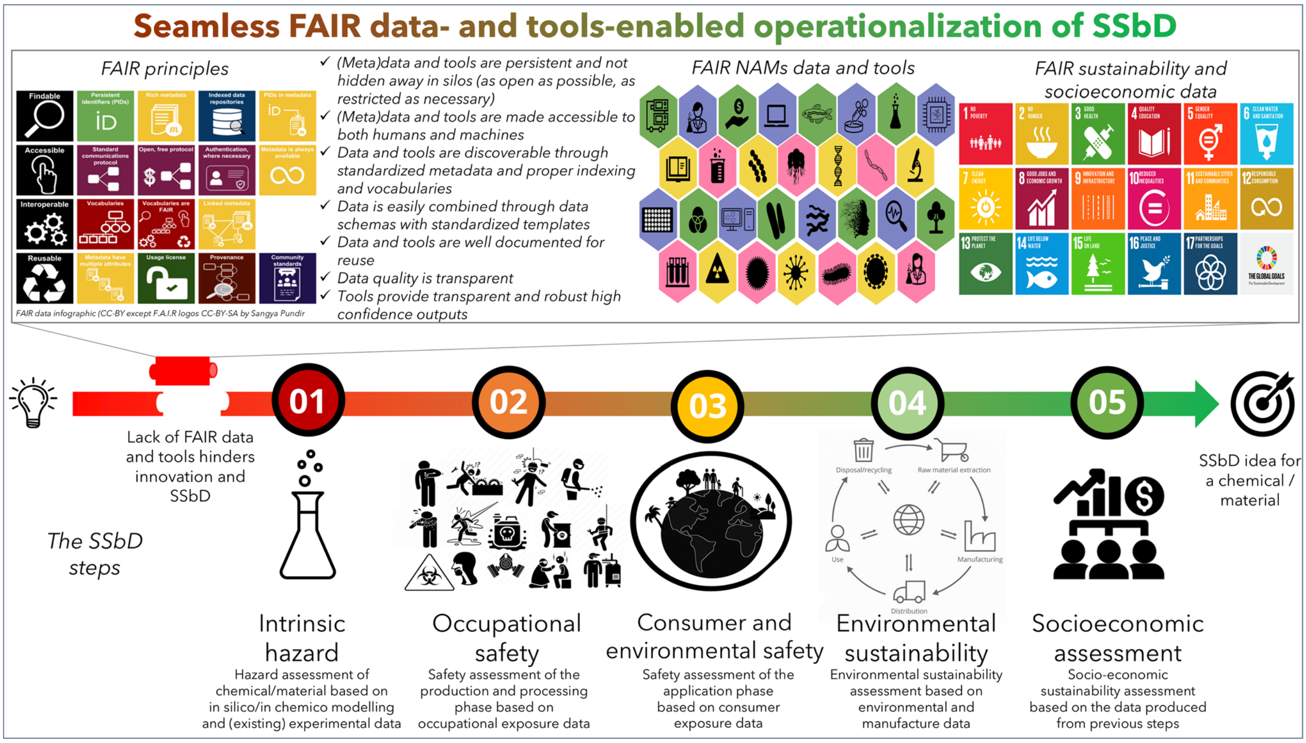

Fig. 2 Depiction of the support that the FAIR principles (upper left) provide during R&I (the early ideation stage is used as an example). The SSbD approach is delineated at the bottom in line with the five steps of the JRC developed framework.[11,25,74]

transparency, quality, and uncertainty) of the integrated data and results generated in the preceding steps. These components are essential for generating a concrete and comparable data-driven SSbD process aimed at achieving human and environmentally friendly chemicals/materials and products. With increasingly FAIR (meta)data and tools, the process can truly become a comprehensive and well-informed outcome for continuously improved decision making.

Ongoing activities to leverage implementation of the FAIR principles into the SSbD approaches

The European PARC project was initiated in 2022 to improve risk assessment of chemicals and materials in order to better protect human health and the environment.[73] Two overarching aims of the project include development and implementation of domain-relevant solutions for FAIRification of next-generation chemical risk assessment data and tools, as well as establishment of a toolbox integrating data and tools to support the JRC SSbD framework.[73,75,76] The computational tools include approaches to assess green, sustainable, and circular chemistry, life cycle assessment (LCA), hazard and risk assessment, and socioeconomic assessment,[11,77] involving methods such as QSARs, other types of NAMs (including *in vitro* data-driven bioinformatics and machine learning), exposure models, cost–benefit analysis, life cycle costing (LCC), and social LCA. Furthermore, models based on AI are emerging to address anticipated data gaps in these data-intense assessments, including *e.g.* AI-driven AOP tools.[78] All subtasks within the SSbD toolbox development align with the FAIR principles. The ultimate objective is the seamless and federated integration of interconnected tools and data, systematically organized into a common functional framework.[73,75,76]

To promote and facilitate data sharing within the SSbD context, particularly within PARC, it is essential to coordinate FAIR e-infrastructures, such as knowledge bases and databases, at a comprehensive level. This includes both data and metadata, as stated by Mech, *et al.*[79] As a result of that, early in the project, the PARC FAIR Data Policy (PFDP)[80] was established. It articulates the guiding principles and stipulations governing data provision, management, access, and reusability within PARC and eventually within the whole chemical/material risk assessment domain.[81] The PFDP[80] is meticulously aligned with the FAIR principles and incorporates due regard for legal considerations, encompassing GDPR compliance, data security, transparency, sustainability, and data quality within the domain of chemical risk assessment. PARC is firmly committed to achieving a high level of data and tool FAIRification, which includes the development of FAIR metadata schemas linked to persistent identifiers as well as increased findability of restricted data through open metadata. To achieve this ambitious endeavor, PARC has adopted the Three Point FAIRification Framework (3PFF) to guide FAIR implementation through the development of domain-relevant metadata requirements (guided by Metadata for Machines, in short M4M workshops), FAIR Implementation Profiles (FIPs) and FAIR Orchestration services (using *e.g.* FAIR Data Points and FAIR Digital Objects, which contribute to a global Internet of FAIR Data and Services) (as described in Magagna, *et al.*,[82] and **https://osf.io/bthf8**). The first two components, *i.e.* domain-relevant metadata schemas and FIPs are particularly relevant in the context of the needed social agreements, as exemplified in Table 1. These systematic approaches to FAIRification, which define metadata requirements, and instances of so-called FAIR-Enabling Resources (findable *via* the search engine FAIR Connect¶¶), respectively, are employed for specific types of data, databases, repositories, and tools, ensure compatibility and harmonization throughout the data and tool ecosystem supporting SSbD. Overall, the initiatives ensure methodically generated standardized metadata schemas aligned with domain-specific standards, formats, and terminologies, as well as enables interoperation between domains, which will be highly useful for practical SSbD operationalization. Finally, PARC proactively explores specialized data repositories and centers, including but not limited to the European Strategy Forum on Research Infrastructures (ESFRI), EIRENE and ELIXIR, as well as MS Open Data initiatives, domain-specific repositories, institutional repositories, and open generic repositories such as Zenodo.[73]

However, due to the scope of PARC, *i.e.* chemical risk assessment, there is currently limited focus on the FAIRness of data relevant for sustainability assessments. Nevertheless, such endeavors can also be envisioned within ongoing and newly started EU projects and other initiatives, including the previously mentioned projects IRISS|||| and HARMLESS***, as well as PINK†††. Thus, a call for communication and discussion across the safety and sustainability dimensions regarding implementation of FAIR principles is strongly suggested.

Conclusions and recommendations

In the present work, we discuss the requisites and advantages associated with the implementation of the FAIR principles within the context of SSbD. Our examination has delved into the surface of how the principles can pave the way for constructing a cutting-edge framework capable of aligning with the ambitious green objectives set forth by the European Commission.[5–8] In the following we provide five overarching recommendations for advancing FAIRification[14] which effectively address some of the current challenges surrounding operationalization of SSbD. The recommendations can be used as a basis for furthering the discussions on the subject within the highly diverse environments affected by the European Commission recommendations on implementation of SSbD.

Recommendation 1

To enable future unanticipated discovery of SSbD-relevant data and tools, we advocate for embedding a broad and harmonized approach to raise the awareness and enable the necessary discussions and agreements regarding *e.g.* the level of persistence for identifiers, minimum information requirements, and (meta)data schemas within the multiple scientific communities that constitute the SSbD domain (Table 1, findability). The outcomes of such scrutiny would allow for the establishment of data quality standards, and delineation of criteria that a dataset should meet to be included among datasets employed for transparent and trustworthy SSbD assessments. It would also support inclusion of harmonized and detailed data documentation, FAQs, and other data descriptions accessible together with the corresponding metadata. Finally, the inclusion of available datasets in search engines would be enhanced as they become increasingly extensively described. Inspiration can be taken from existing FAIR implementation networks, such as the AdvancedNano IN.[44]

Recommendation 2

To effectively support comparable and equal opportunities for performing SSbD across diverse industries, independent of their size, we recommend raising awareness for broad discussions and agreements which are needed regarding the necessary levels of openness of (meta)data (Table 1, accessibility). For example, the need for authentication and/or authorization, as well as the requirements regarding how open (meta)data needs to be in terms of supporting findability, as described above, whilst still retaining necessary levels of restriction in cases where it is inevitable (*e.g.* for data pertaining to GDPR) or desired (*e.g.* for intellectual property reasons), must be agreed upon.

Recommendation 3

For digitalization to truly reach the intended level of interconnectedness and efficiency, we call for the necessary broad discussions and agreements regarding adoption and development of relevant controlled vocabularies and ontologies, as well as qualified references relevant to SSbD, including all relevant dimensions, *i.e.* safety, sustainability and functionality (Table 1, interoperability). These attributes and annotations allow for further unexpected discoveries regarding data capable of supporting diverse assessments in SSbD approaches and makes those approaches flexible to adoption of future developments that we are yet incapable of envisioning. The richness of metadata and the inclusion of references to other (meta)data supports the overall increased interoperability with the global Internet of FAIR Data and Services,[13] allowing for an automated exponential expansion of the SSbD toolbox towards unanticipated coverage. Currently envisioned developments worth mentioning include the previously mentioned Digital Product Passports and solutions for closing chemical/material loops to ensure traceability and enable circularity.[10,83–85]

Recommendation 4

To support the iterative reuse and continuous refinement and improvement of the framework and the whole SSbD domain, yet another aspect of FAIRification includes the discussions and

¶¶ **https://fairconnect.pro/** (accessed December 2023).

|||| **https://iriss-ssbd.eu/** (accessed December 2023).

*** **https://www.harmless-project.eu/** (accessed December 2023).

††† **https://pink-project.eu** (accessed December 2023).

agreements required to maintain continuous documentation, development/inclusion of relevant attributes and licensing policies across relevant databases and repositories (Table 1, reusability). As depicted in Fig. 1, as data builds up, the SSbD process iteratively improves. However, the effect of the iterative refinement can only be realized if data and tools are reliably reusable. As a result, the identification of commonalities among tools is facilitated, opening avenues for discussions on the development of communication protocols between the tools employed within the SSbD community. The integration of embedded models can expedite and simplify this process as well. These models capture semantic relationships and contextual information within words and sentences, enabling a nuanced understanding of word meanings in a multidimensional space for the machines.[86] This approach accelerates the harmonization of tools within the SSbD community. The outlined considerations will bring the scientific community closer to achieving the seamless integration of tools, and because of that SSbD assessors will conduct their analyses without the need to switch data formats between different tools, saving valuable time in the pursuit of the 2050 goals. In this context, the development of APIs for existing tools and databases within the community becomes a crucial focal point. Additionally, the creation of programming libraries (*e.g.*, in R, Java, or Python) utilizing these APIs and converting data from the computational domains of tools into standard data formats significantly enhances interoperability.

Here, it becomes relevant to recommend an overall focus on approaches employed to harmonize FAIRification through the development of *e.g.* (meta)data schemas and FIPs, such as the approach taken within PARC (see details above). These harmonization efforts are particularly important within broad communities where multiple data types are shared from numerous sources, such as the R&I (and hence, SSbD) domain. Harmonized (meta)data schemas and implementation of FAIR-Enabling Resources, in line with FIPs, ensure seamless integration of (meta)data and tools.[13]

Recommendation 5

Finally, a significant challenge to overcome and thereby a clear recommendation is to not forget the necessary discussions on provenance of data and services (Table 1, reusability), *i.e.* how data is collected/developed, by whom, the intended use, and the challenges it aims to address. Technological solutions, such as blockchain technologies, have been suggested to address such needs. However, the establishment of licenses governing data/tool use and reuse, such as the Creative Commons licenses, is a critical step forward in this direction. Streamlining the procedures for issuing these licenses, which are often time-consuming, is equally imperative. Furthermore, the active involvement of industry is paramount, necessitating the contribution of data and practical experience. Achieving this requires a substantial enhancement in the reliability of data transfer methods, not only among SSbD assessors but also between the different tools used during the assessment. Highly confidential data could be handled through use of federated tools which allow for processing without downloading.[42,87] Building credibility within the stakeholders is crucial to encourage practical contributions to the community.

Overall, digitalization, reuse of existing data, effective use of new scientific knowledge/developments (*e.g.* application of NAMs), transparency and trust, and seamlessness, all depend on implementation of FAIRification procedures at all levels and in all aspects of SSbD operationalization; from assessment of hazard and risk to the broad sustainability and socioeconomic impacts of chemicals/materials. However, only by stepping up and acknowledging the efforts needed for FAIRification, can it become reality, supporting development of seamlessly connected FAIR data and tools that automatically allow for comparable, transparent, and trustworthy assessments and predictions effectively supporting well-informed decisions capable of avoiding unbalanced trade-offs between safety, sustainability, and functionality. Thereby, allowing the pre-market approach, SSbD, to lead us towards more planet-centric and forward-looking systems and business models focused on rethinking, restoring and replenishing, instead of the traditional thinking around "merely" reducing and recycling.[10]

Data availability

No new data were created or analysed in this study. Thus, data sharing is not applicable to this article.

Author contributions

Achilleas Karakoltzidis: conceptualization, investigation, writing – original draft, writing – review & editing, visualization. Chiara Laura Battistelli: conceptualization, investigation, writing – original draft, writing – review & editing. Cecilia Bossa: conceptualization, investigation, writing – original draft, writing – review & editing. Evert A. Bouman: conceptualization, investigation, writing – original draft, writing – review & editing. Irantzu Garmendia Aguirre: writing – review & editing. Maryam Zare Jeddi: conceptualization, investigation, writing – original draft, writing – review & editing. Spyros Karakitsios: visualization, writing – review & editing. Magnus Løfstedt: conceptualization, investigation, writing – original draft, writing – review & editing. Barbara Magagna: writing – review & editing. Ivo Iavicoli: writing – review & editing, funding acquisition. Veruscka Leso: writing – review & editing. Denis Sarigiannis: writing – review & editing, funding acquisition. Erik Schultes: writing – review & editing. Lya G. Soeteman-Hernández: conceptualization, investigation, visualization, writing – original draft, writing – review & editing. Vrishali Subramanian: conceptualization, investigation, project administration, writing – original draft, writing – review & editing. Penny Nymark: conceptualization, investigation, writing – original draft, writing – review & editing, visualization, supervision, funding acquisition.

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgements

This study is part of the Partnership for the Assessment of Risk from Chemicals (PARC), funded by the European Union's Horizon Europe Research and Innovation Programme under grant agreement no. 101057014. The study was also supported by Horizon Europe projects PINK (grant agreement no. 101137809) and INSIGHT (grant agreement no. 101137742). The authors express their gratitude to Dr Martine Bakker, Dr Jaco Westra, and Dr Jacqueline van Engelen from the National Institute for Public Health and the Environment (RIVM) for their valuable review and feedback on this paper.

References

1 M. Antikainen, T. Uusitalo and P. Kivikytö-Reponen, *Procedia CIRP*, 2018, **73**, 45–49, DOI: 10.1016/j.procir.2018.04.027.

2 M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos and P. E. Bourne, *Sci. Data*, 2016, **3**, 1–9, DOI: 10.1038/sdata.2016.18.

3 M. Barker, N. P. Chue Hong, D. S. Katz, A.-L. Lamprecht, C. Martinez-Ortiz, F. Psomopoulos, J. Harrow, L. J. Castro, M. Gruenpeter and P. A. Martinez, *Sci. Data*, 2022, **9**, 622, DOI: 10.1038/s41597-022-01710-x.

4 T. Hartung, *Front. Artif. Intell.*, 2023, **6**, 1269932, DOI: 10.3389/frai.2023.1269932.

5 European Commission, *EU's Chemicals Strategy for Sustainability*, European Commission, 2020, https://ec.europa.eu/environment/pdf/chemicals/2020/10/Strategy.pdf.

6 European Commission, *Document 52018DC0773*, European Commission, Brussels, 2018, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0773.

7 European Commission, 2021, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021DC0400&qid=1623311742827.

8 European Commission, 2020, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2020%3A667%3AFIN.

9 B. Suárez-Eiroa, E. Fernández and G. Méndez, *J. Clean. Prod.*, 2021, **322**, 129071, DOI: 10.1016/j.jclepro.2021.129071.

10 C. Apel, K. Kümmerer, A. Sudheshwar, B. Nowack, C. Som, C. Colin, L. Walter, J. Breukelaar, M. Meeus and B. Ildefonso, *Curr. Opin. Green Sustainable Chem.*, 2023, 100876, DOI: 10.1016/j.cogsc.2023.100876.

11 European Commission: Joint Research Centre, C. Caldeira, L. Farcal I. Garmendia Aguirre, L. Mancini, *et al.*, *Safe and sustainable by design chemicals and materials – Framework for the definition of criteria and evaluation procedure for chemicals and materials*, Publications Office of the European Union, 2022, https://data.europa.eu/doi/10.2760/487955.

12 P. Nymark, M. Bakker, S. Dekkers, R. Franken, W. Fransman, A. García-Bilbao, D. Greco, M. Gulumian, N. Hadrup and S. Halappanavar, *Small*, 2020, **16**, 1904749, DOI: 10.1002/smll.201904749.

13 E. Schultes, *FAIR Connect*, 2023, **1**, 13–17, DOI: 10.3233/FC-221514.

14 E. Huerta, B. Blaiszik, L. C. Brinson, K. E. Bouchard, D. Diaz, C. Doglioni, J. M. Duarte, M. Emani, I. Foster and G. Fox, *Sci. Data*, 2023, **10**, 487, DOI: 10.1038/s41597-023-02298-6.

15 A. Landi, M. Thompson, V. Giannuzzi, F. Bonifazi, I. Labastida, L. O. B. d. S. Santos and M. Roos, *Data Intelligence*, 2020, **2**, 47–55, DOI: 10.1162/dint_a_00027.

16 P. Fantke, C. Cinquemani, P. Yaseneva, J. De Mello, H. Schwabe, B. Ebeling and A. A. Lapkin, *Chem*, 2021, **7**, 2866–2882, DOI: 10.1016/j.chempr.2021.09.012.

17 C. Ding, C. Liu, C. Zheng and F. Li, *Sustainability*, 2021, **14**, 216, DOI: 10.3390/su14010216.

18 J. Han, N. Lin, J. Ruan, X. Wang, W. Wei and H. Lu, *IEEE Internet Things J.*, 2020, **8**, 9683–9696, DOI: 10.1109/JIOT.2020.3037729.

19 Y. Lin, J. Yang, Z. Lv, W. Wei and H. Song, *Sensors*, 2015, **15**, 20925–20944, DOI: 10.3390/s150820925.

20 C. Zhang and Y. Lu, *J. Ind. Inf. Integr.*, 2021, **23**, 100224, DOI: 10.1016/j.jii.2021.100224.

21 S.-W. Chen, D. L. Chiang, C.-H. Liu, T.-S. Chen, F. Lai, H. Wang and W. Wei, *J. Med. Syst.*, 2016, **40**, 1–12, DOI: 10.1007/s10916-016-0484-7.

22 Y. He, J. He and N. Wen, *J. Innov. Knowl.*, 2023, **8**, 100347, DOI: 10.1016/j.future.2018.05.080.

23 L. Cai, Y. Qi, W. Wei, J. Wu and J. Li, *Future Generat. Comput. Syst.*, 2019, **93**, 570–582, DOI: 10.1016/j.future.2018.05.080.

24 L. Yang, H. Zou, C. Shang, X. Ye and P. Rani, *Technol. Forecast. Soc. Change*, 2023, **188**, 122308, DOI: 10.1016/j.techfore.2022.122308.

25 European Commission: Joint Research Centre, C. Caldeira, I. Garmendia Aguirre, D. Tosches, L. Mancini, *et al.*, *Safe and Sustainable by Design Chemicals and Materials – Application of the SSbD Framework to Case Studies*, Publications Office of the European Union, 2023, https://data.europa.eu/doi/10.2760/329423.

26 N. Kleinstreuer and T. Hartung, *Arch. Toxicol.*, 2024, 1–20, DOI: 10.1007/s00204-023-03666-2.

27 Y. Jiang, S. Yin, K. Li, H. Luo and O. Kaynak, *Philos. Trans. R. Soc., A*, 2021, **379**, 20200360, DOI: 10.1098/rsta.2020.0360.

28 J. Leng, D. Wang, W. Shen, X. Li, Q. Liu and X. Chen, *J. Manuf. Syst.*, 2021, **60**, 119–137, DOI: 10.1016/j.jmsy.2021.05.011.

29 European Commission, 2020, https://www.consilium.europa.eu/en/press/press-releases/2020/12/17/council-approves-conclusions-on-making-the-recovery-circular-and-green/.

30 J. T. K. Quik, J. P. A. Lijzen and J. Spijker, *Creating safe and sustainable material loops in a circular economy: proposal for a tiered modular framework to assess options for material recycling*, RIVM Report 2018-0173, RIVM Official Reports, 2019, DOI: 10.21945/RIVM-2018-0173.

31 C. Wittwehr, L.-A. Clerbaux, S. Edwards, M. Angrish, H. Mortensen, A. Carusi, M. Gromelski, E. Lekka,

V. Virvilis, M. Martens, *et al.*, *ALTEX*, 2024, **41**(1), 50–56, DOI: 10.14573/altex.2307131.

32 European Commission, *Proposal for a regulation establishing a common data platform on chemicals, laying down rules to ensure that the data contained in it are findable, accessible, interoperable and reusable and establishing a monitoring and outlook framework for chemicals*, COM(2023)779, 2023, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52023PC0779.

33 W. Huang, K. Y. Chau, I. Y. Kit, N. Nureen, M. Irfan and A. Dilanchiev, *Front. Psychol.*, 2022, **13**, 930138, DOI: 10.3389/fpsyg.2022.930138.

34 G. Li, R. Zhang, S. Feng and Y. Wang, *Bus. Strat. Environ.*, 2022, **31**, 3574–3594, DOI: 10.1002/bse.3105.

35 V. Alcácer and V. Cruz-Machado, *Eng. Sci. Technol.*, 2019, **22**, 899–919, DOI: 10.1016/j.jestch.2019.01.006.

36 A. Bharadwaj, O. A. El Sawy, P. A. Pavlou and N. V. Venkatraman, *MIS Quart.*, 2013, **37**(2), 471–482.

37 A. Ghezzi, A. Cavallaro, A. Rangone and R. Balocco, *Int. J. Technol. Manag.*, 2015, **68**, 21–48, DOI: 10.1504/IJTM.2015.068777.

38 T. Zonta, C. A. Da Costa, R. da Rosa Righi, M. J. de Lima, E. S. da Trindade and G. P. Li, *Comput. Ind. Eng.*, 2020, **150**, 106889, DOI: 10.1016/j.cie.2020.106889.

39 I. Furxhi, A. Costa, S. Vázquez-Campos, C. Fito-López, D. Hristozov, J. A. T. Ramos, S. Resch, M. Cioffi, S. Friedrichs and C. Rocca, *RSC Sustainability*, 2023, **1**, 234–250, DOI: 10.1039/D2SU00101B.

40 I. Furxhi, *NanoImpact*, 2022, **25**, 100378, DOI: 10.1016/j.impact.2021.100378.

41 C. Bossa, C. Andreoli, M. Bakker, F. Barone, I. De Angelis, N. Jeliazkova, P. Nymark and C. L. Battistelli, *Comput. Toxicol.*, 2021, **20**, 100190, DOI: 10.1016/j.comtox.2021.100190.

42 N. Jeliazkova, M. D. Apostolova, C. Andreoli, F. Barone, A. Barrick, C. Battistelli, C. Bossa, A. Botea-Petcu, A. Châtel and I. De Angelis, *Nat. Nanotechnol.*, 2021, **16**, 644–654, DOI: 10.1038/s41565-021-00911-6.

43 S. Watford, S. Edwards, M. Angrish, R. S. Judson and K. P. Friedman, *Toxicol. Appl. Pharmacol.*, 2019, **380**, 114707, DOI: 10.1016/j.taap.2019.114707.

44 V. I. Dumit, A. Ammar, M. I. Bakker, M. A. Bañares, C. Bossa, A. Costa, H. Cowie, D. Drobne, T. E. Exner and L. Farcal, *Nano Today*, 2023, **51**, 101923, DOI: 10.1016/j.nantod.2023.101923.

45 European Commission, *European Research Data Landscape – Final Report*, Publications Office of the European Union, 2022, DOI: 10.2777/3648, https://zenodo.org/communities/erdl21/?page=1&size=20.

46 M. D. Wilkinson, M. Dumontier, S.-A. Sansone, L. O. Bonino da Silva Santos, M. Prieto, D. Batista, P. McQuilton, T. Kuhn, P. Rocca-Serra and M. Crosas, *Sci. Data*, 2019, **6**, 174, DOI: 10.1038/s41597-019-0184-5.

47 A. Ammar, S. Bonaretti, L. Winckers, J. Quik, M. Bakker, D. Maier, I. Lynch, J. van Rijn and E. Willighagen, *Nanomaterials*, 2020, **10**, 2068, DOI: 10.3390/nano10102068.

48 N. Krans, A. Ammar, P. Nymark, E. Willighagen, M. Bakker and J. Quik, *NanoImpact*, 2022, **27**, 100402.

49 FAIR Data Maturity Model Working Group, *FAIR Data Maturity Model. Specification and Guidelines (1.0)*, Zenodo, 2020, DOI: 10.15497/rda00050.

50 F. Belliard, A. M. Maineri, E. Plomp, A. F. Ramos Padilla, J. Sun and M. Zare Jeddi, *PLoS Comput. Biol.*, 2023, **19**, e1011668, DOI: 10.1371/journal.pcbi.1011668.

51 D. C. Berrios, A. Beheshti and S. V. Costes, FAIRness and Usability for Open-Access Omics Data Systems, *AMIA Annu. Symp. Proc.*, 2018, 232–241, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6371294/.

52 EPA, 2018, https://www.epa.gov/sites/default/files/2018-06/documents/epa_alt_strat_plan_6-20-18_clean_final.pdf.

53 ECHA, *New Approach Methodologies in Regulatory Science*, 2016, https://op.europa.eu/en/publication-detail/-/publication/4c2ad7eb-9ad0-11e6-868c-01aa75ed71a1/language-en.

54 A. O. Stucki, T. S. Barton-Maclaren, Y. Bhuller, J. E. Henriquez, T. R. Henry, C. Hirn, J. Miller-Holt, E. G. Nagy, M. M. Perron and D. E. Ratzlaff, *Front. Toxicol.*, 2022, **4**, 964553, DOI: 10.3389/ftox.2022.964553.

55 S. Schmeisser, A. Miccoli, M. von Bergen, E. Berggren, A. Braeuning, W. Busch, C. Desaintes, A. Gourmelon, R. Grafström and J. Harrill, *Environ. Int.*, 2023, 108082, DOI: 10.1016/j.envint.2023.108082.

56 OECD, *Guideline No. 497: Defined Approaches on Skin Sensitisation*, OECD Publishing, 2021, https://www.oecd-ilibrary.org/deliver/b92879a4-en.pdf?itemId=%2Fcontent%2Fpublication%2Fb92879a4-en&mimeType=pdf.

57 M. Alimohammadi, B. Meyburg, A. K. Ückert, A. K. Holzer and M. Leist, *EFSA Support. Publ.*, 2023, **20**, 7794E, DOI: 10.2903/sp.efsa.2023.EN-7793/full.

58 F. R. Cassee, E. A. Bleeker, C. Durand, T. Exner, A. Falk, S. Friedrichs, E. Heunisch, M. Himly, S. Hofer and N. Hofstätter, *Comput. Struct. Biotechnol. J.*, 2024, **25**, 105–126, DOI: 10.1016/j.csbj.2024.05.018.

59 E. Carnesecchi, I. Langezaal, P. Browne, S. Batista-Leite, I. Campia, S. Coecke, B. Dagallier, P. Deceuninck, J. L. C. Dorne and J. V. Tarazona, *Regul. Toxicol. Pharmacol.*, 2023, 105426, DOI: 10.1016/j.yrtph.2023.105426.

60 K. Alberi, M. B. Nardelli, A. Zakutayev, L. Mitas, S. Curtarolo, A. Jain, *et al.*, *J. Phys. D: Appl. Phys.*, 2019, **52**(1), 013001, DOI: 10.1088/1361-6463/aad926.

61 H.-J. Klimisch, M. Andreae and U. Tillmann, *Regul. Toxicol. Pharmacol.*, 1997, **25**, 1–5, DOI: 10.1006/rtph.1996.1076.

62 OECD, *Guideline No. 386: (Q)SAR Assessment Framework: Guidance for the Regulatory Assessment of (Quantitative) Structure–Activity Relationship Models, Predictions, and Results based on Multiple Predictions*, OECD Publishing, Paris, 2023, https://www.oecd.org/chemicalsafety/risk-assessment/qsar-assessment-framework.pdf.

63 N. Roth, J. Zilliacus and A. Beronius, *Front. Toxicol.*, 2021, **3**, 746430, DOI: 10.3389/ftox.2021.746430.

64 A. Beronius, L. Molander, C. Rudén and A. Hanberg, *J. Appl. Toxicol.*, 2014, **34**, 607–617, DOI: 10.1002/jat.2991.

65 G. Shao, A. Beronius and P. Nymark, *Front. Toxicol.*, 2023, **5**, 1319985, DOI: 10.3389/ftox.2023.1319985.
66 I. Furxhi, R. Bengalli, G. Motta, P. Mantecca, O. Kose, M. Carriere, E. U. Haq, C. O'mahony, M. Blosi and D. Gardini, *ACS Appl. Nano Mater.*, 2023, **6**, 3948–3962.
67 M. T. Cronin, S. J. Belfield, K. A. Briggs, S. J. Enoch, J. W. Firman, M. Frericks, C. Garrard, P. H. Maccallum, J. C. Madden and M. Pastor, *Regul. Toxicol. Pharmacol.*, 2023, **140**, 105385, DOI: 10.1016/j.yrtph.2023.105385.
68 N. Ravi, P. Chaturvedi, E. Huerta, Z. Liu, R. Chard, A. Scourtas, K. Schmidt, K. Chard, B. Blaiszik and I. Foster, *Sci. Data*, 2022, **9**, 657, DOI: 10.1038/s41597-022-01712-9.
69 C. Rivetti and B. Campos, *Integr. Environ. Assess. Manage.*, 2023, **19**, 571–573, DOI: 10.1002/ieam.4763.
70 L. Maltby, R. Brown, J. H. Faber, N. Galic, P. J. Van den Brink, O. Warwick and S. Marshall, *Sci. Total Environ.*, 2021, **791**, 148631, DOI: 10.1016/j.scitotenv.2021.148631.
71 M. T. Cronin, J. C. Madden, C. Yang and A. P. Worth, *Comput. Toxicol.*, 2019, **10**, 38–43, DOI: 10.1016/j.comtox.2018.12.006.
72 M. J. Moné, G. Pallocca, S. E. Escher, T. Exner, M. Herzler, S. H. Bennekou, H. Kamp, E. D. Kroese, M. Leist and T. Steger-Hartmann, *Arch. Toxicol.*, 2020, **94**, 3581–3592, DOI: 10.1007/s00204-020-02866-4.
73 P. Marx-Stoelting, G. Rivière, M. Luijten, K. Aiello-Holden, N. Bandow, K. Baken, A. Cañas, A. Castano, S. Denys, C. Fillol, M. Herzler, I. Iavicoli, S. Karakitsios, J. Klanova, M. Kolossa-Gehring, A. Koutsodimou, J. Lobo Vicente, I. Lynch, S. Namorado, S. Norager, A. Pittman, S. Rotter, D. Sarigiannis, M. J. Silva, J. Theunis, T. Tralau, M. Uhl, J. van Klaveren, L. Wendt-Rasch, E. Westerholm, C. Rousselle and P. Sanders, *Arch. Toxicol.*, 2023, 1–16, DOI: 10.1007/s00204-022-03435-7.
74 GDPR, *Regulation (EU)*, 2016, vol. 679, p. 2016.
75 D. Sarigiannis, F. Nikiforou, A. Karakoltzidis, A. Gypakis and S. Karakitsios, *AICHE Conference 2023*, 2023.
76 D. Sarigiannis, F. Nikiforou, A. Karakoltzidis, A. Gypakis and S. Karakitsios, *Toxicol. Lett.*, 2023, **384**(1), S184–S185, DOI: 10.1016/S0378-4274(23)00692-6.
77 V. Subramanian, W. J. Peijnenburg, M. G. Vijver, C. F. Blanco, S. Cucurachi and J. B. Guinée, *Chemosphere*, 2022, 137080, DOI: 10.1016/j.chemosphere.2022.137080.
78 S. Kumar, N. ZaidKilani, S. Khalid, L. Slater, A. Karakoltzidis, N. Papaioannou, P. Nymark, K. E. Tollefsen, H. A. Rashwan, K. Audouze and V. Kumar, *Toxicol. Lett.*, 2023, **384**, S105, DOI: 10.1016/S0378-4274(23)00513-1.
79 A. Mech, S. Gottardo, V. Amenta, A. Amodio, S. Belz, S. Bøwadt, J. Drbohlavová, L. Farcal, P. Jantunen, A. Małyska, K. Rasmussen, J. Riego Sintes and H. Rauscher, *Regul. Toxicol. Pharmacol.*, 2022, **128**, 105093, DOI: 10.1016/j.yrtph.2021.105093.
80 R. Stierum, S. Fraize-Frontier, O. Frezot, B. Magagna, E. Schultes, C. Bossa, P. Nymark, T. Pronk, K. Řiháčková, L. Bielska, R. Hulek, J. Bouwman, S. Le Dévédec, I. Lynch, O. Májek, J. Theunis, S. Remy, P. Von der Ohe, I. Huskova and F. van de Brug, *Deliverable D7.1: PARC Data Management Plan V1.0 (DMP)*, 2022, https://www.eu-parc.eu/deliverables.
81 F. Brug, R. Stierum, S. Fraize-Frontier, O. Frezot, B. Magagna, E. Schultes, C. Bossa, P. Nymark, T. Pronk, K. Řiháčková, L. Bielska, R. Hulek, J. Bouwman, S. Dévédec, I. Lynch, O. Májek, J. Theunis, S. Remy, P. Ohe and I. Huskova, *Deliverable D7.1 PARC Data Management Plan V1.0 (DMP) WP 7.1*, 2023, https://www.eu-parc.eu/sites/default/files/2023-10/PARC_D7.1.pdf.
82 B. Magagna, E. Schultes, M. Suchánek and T. Kuhn, *Res. Ideas Outcomes*, 2022, **8**, e94451, DOI: 10.3897/rio.8.e94451.
83 T. Adisorn, L. Tholen and T. Götz, *Energies*, 2021, **14**, 2289, DOI: 10.3390/en14082289.
84 S. F. Jensen, J. H. Kristensen, S. Adamsen, A. Christensen and B. V. Waehrens, *Sustain. Prod. Consum.*, 2023, **37**, 242–255, DOI: 10.1016/j.spc.2023.02.021.
85 D. J. Langley, E. Rosco, M. Angelopoulos, O. Kamminga and C. Hooijer, *J. Bus. Res.*, 2023, **169**, 114259, DOI: 10.1016/j.jbusres.2023.114259.
86 D. Bianchini, V. De Antonellis and M. Garda, *Knowl. Inf. Syst.*, 2024, **66**, 1469–1502, DOI: 10.1007/s10115-023-02014-1.
87 J. Simm, L. Humbeck, A. Zalewski, N. Sturm, W. Heyndrickx, Y. Moreau, B. Beck and A. Schuffenhauer, *J. Cheminf.*, 2021, **13**, 1–14, DOI: 10.1186/s13321-021-00576-2.