TRANSIENCE

TRANSITIONING TOWARDS AN EFFICIENT, CARBON-NEUTRAL CIRCULAR EUROPEAN INDUSTRY

Date: 25/06/2025

D3.8 – Open science protocols

WP3 – Characterising circularity and decarbonisation opportunities – generating model inputs

Disclaimer

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Health and Digital Executive Agency (HaDEA). Neither the European Union nor the granting authority can be held responsible for them.

Copyright Message

This report, if not confidential, is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0); a copy is available here: https://creativecommons.org/licenses/by/4.0/. You are free to share (copy and redistribute the material in any medium or format) and adapt (remix, transform, and build upon the material for any purpose, even commercially) under the following terms: (i) attribution (you must give appropriate credit, provide a link to the license, and indicate if changes were made; you may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use); (ii) no additional restrictions (you may not apply legal terms or technological measures that legally restrict others from doing anything the license permits).

Grant Agreement Number	101137606	**Acronym**	TRANSIENCE
Full Title	TRANSItioning towards an Efficient, carbon-Neutral Circular European industry		
Topic	HORIZON-CL4-2023-TWIN-TRANSITION-01-36		
Funding scheme	HORIZON EUROPE, RIA – Research and Innovation Action		
Start Date	January 2024	**Duration**	48 Months
Project URL	https://www.transience.eu/		
EU Project Advisor	Fatima GONZALEZ COMEZ		
Project Coordinator	Institute of Communication and Computer Systems (ICCS)		
Deliverable	D3.8 – Open science protocols		
Work Package	WP3 – Characterising circularity and decarbonisation opportunities – generating model inputs		
Date of Delivery	**Contractual** 30/06/2025	**Actual**	25/06/2025
Nature	Report	**Dissemination Level**	Public
Lead Beneficiary	HOLISTIC IKE (HOL)		
Responsible Authors	Georgios Xexakis [HOL], Stratis Alexandrou [ICCS]		
Contributors	Konstantinos Koasidis [ICCS]		
Reviewer(s)	Christina Tigka, Alexandros Nikas [ICCS]; Panagiotis Fragkos [E3M]; Jakob Duerrwaechter [PIK]; Gergő Sütő [UU]; Alvaro Calzadilla [UCL]		
Keywords	open source; open access; data management; data exchange; nomenclature.		

EC Summary Requirements

1. Changes with respect to the Description of the Action (DoA)

No changes with respect to the work described in the DoA.

2. Dissemination and uptake

The deliverable will be primarily used by the TRANSIENCE partners as it provides them with guidelines and protocols for open access, as well as associated checklists and ontologies to be employed in the second phase of the project, where all individual modules will be integrated to the MIC3 model. This information can also be useful for external audiences of the project (e.g., academics, industries) to understand the type of data that will be used in the modules.

3. Short summary of results (<250 words)

Open science is a set of principles and practices to make scientific research available and accessible to everyone. This deliverable aims to explore and operationalise open science principles and practices that can support the research activities of the TRANSIENCE project. First, the report summarises the main activities and outputs of TRANSIENCE and matches them with relevant open science principles. Specifically, all research activities of the project will be described through open-access publications (deliverable reports, scientific journal papers, etc.). All input and output datasets developed in TRANSIENCE will be shared based on the FAIR principles for enabling findable, accessible, interoperable, and reusable data. Model code as well as relevant scripts (e.g., for data processing) will be shared as open source, while both code and data will be deposited in online repositories that follow the TRUST (Transparency, Responsibility, User focus, Sustainability, and Technology) principles. Since modelling activities play a major role in TRANSIENCE, the report provides more details on open science practices that can support modelling. Specifically, the report presents open science protocols to guide major research activities related to model development, linking, and use. Subsequently, the report delves further into practices that support the exchange of model data by introducing the concept of ontologies and providing a hands-on example of applying ontologies in the modelling data of TRANSIENCE. The open science protocols and ontologies will be further updated later in the project based on the feedback and experience of the modelling partners.

4. Evidence of accomplishment

This report and the accompanying GitHub repository[1] containing the ontologies and nomenclature of input data.

[1] https://github.com/i2amparis/transience_open_science

Preface

The need to approach climate action, resource efficiency, and circularity performance as integrated, economy-wide, cross-cutting issues is growingly gaining attention in the policy world, stimulating the development of new industrial policies in Europe and worldwide. Currently, however, there is little progress in conceptualising the circular economy and understanding its interactions with climate action. State-of-the-art modelling capacity to capture the interplay of the two agendas and their implications for energy-intensive sectors as well as to represent the European industry's transformation in line with the region's vision for climate neutrality is not yet fully developed. TRANSIENCE will undertake a comprehensive characterisation and assessment of circularity principles and measures vis-à-vis decarbonisation, by looking at the twin transition of European industries through the lenses of global competitiveness, innovation, and holistic sustainability. It will then produce MIC3, a consistent, fully open-source model ecosystem to assess industrial circularity, decarbonisation, and sustainability. A series of interoperable modules on the socioeconomic, service and product, material, industrial, energy-system, and environmental perspectives of the transformation of European industry will be developed and integrated, building on and opening the code of leading modelling tools. MIC3 will finally be used in extensive scenario modelling to produce diverse pathways toward a material-efficient, circular, climate-neutral, sustainable European industry. Transparency, openness, and knowledge sharing will be promoted, and technical capacities will be developed in four industrial agglomerations in the EU, moving beyond stakeholder consultation, onto model co-development, continuous validation of assumptions, co-creation of scenario modelling, evaluation of the desirability and usability of the developed model and insights, and eventually co-production of science and action.

ICCS – Institute of Communication and Computer Systems	EL
CEPS – Centre for European Policy Studies	BE
E3M – E3-Modelling AE	EL
Fraunhofer – Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.	DE
HOL – HOLISTIC IKE	EL
PIK – Potsdam Institut für Klimafolgenforschung e.V.	DE
PNTEC – Park Naukowo-Technologiczny Euro-Centrum Spolka Z Ograniczona Odpowiedzialnoscia	PL
TECNALIA – Fundación Tecnalia Research & Innovation	ES
UU – Universiteit Utrecht	NL
WI – Wuppertal Institut für Klima, Umwelt, Energie gGmbH	DE
PSI – Paul Scherrer Institut	CH
UCL – University College London	UK

Funded by the European Union

Executive Summary

Open science is a set of principles and practices to make scientific research available and accessible to everyone. This report aims to explore and operationalise open science principles and practices that can support the research activities of the TRANSIENCE project. The report is part of Task 3.1, "Open and FAIR data management (Phase 1)", which aims to lay the groundwork and support knowledge exchange in the project.

First, the report summarises the main activities and outputs of TRANSIENCE and matches them with relevant open science principles. Specifically, all research activities of the project will be described through open-access publications (deliverable reports, journal papers, etc.) while their input and output datasets will be shared based on the FAIR principles for enabling findable, accessible, interoperable, and reusable data. Model code as well as relevant scripts (e.g., for data processing) will be shared as open source, while both code and data will be deposited in online repositories that follow the TRUST principles (Transparency, Responsibility, User focus, Sustainability, and Technology).

Since modelling activities play a major role in TRANSIENCE, deliverable D3.8 provides more details on open science practices that can support modelling. Specifically, the deliverable presents open science protocols to guide major research activities related to model development, linking, and use. The protocols have been inspired by the open science protocols created in the Horizon Europe IAM COMPACT[2] project and were informed by studies that synthesised and enhanced recommendations from the literature in terms of improving the development and linking of integrated assessment models (IAMs), energy system models, environmental models, etc.

Examples of relevant practices for model development include sharing the model code and updates in GitHub, sharing data through the TRANSIENCE community in Zenodo, and documenting the model in a Read the Docs page. Similarly, the model linking protocol includes recommendations such as documenting the harmonisation of input data and sharing both processing code and resulting datasets in platforms supporting open science practices. Suggestions for model use include an interactive tool for validating modelling data and the use of the IAM PARIS platform for visualising all scenario results. Subsequently, the deliverable delves further into practices that support the exchange of model data by introducing the concept of ontologies and providing a hands-on example of applying ontologies in the modelling data of TRANSIENCE.

The open science protocols and ontologies will be further updated later in the project based on the feedback and experience of the modelling partners. As model linking will be critical for creating the interconnected MIC3 model, the protocols will be enriched with more detailed information about the development of application programming interfaces (APIs) to facilitate these links and the packaging of models and their system requirements in Docker images. The ontology will also be further expanded to include all input and output data from the individual models, facilitating the mapping of the different data transfers that need to take place in TRANSIENCE. Finally, we will aim to standardise and disseminate the protocols and ontologies as much as possible to make them useful and usable by modellers beyond TRANSIENCE, e.g., by publishing them in scientific publications and data repositories.

[2] https://www.iam-compact.eu

Contents

Table of Figures

Table of Tables

1 Introduction

Open science is a set of principles and practices to make scientific research available and accessible to everyone (UNESCO, 2022). As the name implies, the goal of open science is to make scientific knowledge as open as possible, by providing unrestricted access to scientific publications and research data (open-access publishing), by making software and its source code available (open-source code), and by providing open resources to support scientific uptake such as open educational material or even hardware. Nevertheless, open science is not only about opening up scientific outputs, but also about involving different stakeholders throughout the scientific process (UNESCO, 2022). As Figure 1 shows, open science includes efforts to develop a dialogue with other knowledge systems (e.g., local communities, indigenous people) in a transdisciplinary manner, engage societal actors (e.g., through citizen and participatory science), and provide access to virtual and physical science infrastructure (e.g., open bibliometric systems). While open science aims to reduce the barriers and restrictions on the availability of scientific knowledge, it does not necessarily mean cost-free or even uncontrolled sharing (UNESCO, 2022). Many open science practices rely on licensing, defining clear conditions for the users such as providing appropriate credit to publication authors. Additionally, there are many open science initiatives that aim to make science cost free by providing free publication platforms or covering the cost of publishing for researchers.

Figure 1. The four main pillars of open science, reproduced from UNESCO (2022)

Funded by the European Union

This report, in the context of deliverable D3.8 'Open science protocols', aims to explore and operationalise open science principles and practices that will be relevant to the research activities of the TRANSIENCE project. The report is part of Task 3.1 'Open and FAIR data management (Phase 1)', that aims to lay the groundwork and support knowledge exchange in the project. The first part of this task was to build a basis for this exchange by developing an open, machine-actionable Data Management Plan (DMP)[3] in the context of deliverable D3.1 'Machine-actionable, open data management plan'. While the DMP touched upon some aspects of open science, such as ensuring that project data are findable, accessible, interoperable, and reusable (FAIR), D3.8 aims to provide more detailed and hands-on guidelines for achieving open science in TRANSIENCE, especially in terms of its modelling activities. On the latter, the report also complements the open model development strategy[4] that has been created during the first project year and which aimed to synthesise insights from the stakeholder activities of the project in a list of required model features.

First, this report summarises the main activities and outputs of TRANSIENCE and matches them with relevant open science principles (Section 2). Then, it elaborates on these principles and proposes ways to follow them in TRANSIENCE, for instance by sharing open-access datasets on the open-science platform Zenodo[5] and open-source code on GitHub[6] repositories. Since modelling activities play a major role in TRANSIENCE, Section 3 and 4 provide more details on open-science practices that can support modelling development activities. Specifically, Section 3 presents open-science protocols to guide major research activities related to model development, model linking, and model use. Section 4 delves further into practices that support the exchange of model data by introducing the concept of ontologies and providing a hands-on example of applying ontologies in modelling data in TRANSIENCE. Both protocols and ontologies will be further updated as the project's modelling activities progress. After enhancing and opening the individual models in Phase 1 of TRANSIENCE, Phases 2 and 3 include testing these models, establishing interconnections between them to form the MIC3 model, and using both the individual and linked models in four case studies for the European industry and other modelling exercises for the EU and its Member States. More details about next steps and future updates of the protocols and ontologies are provided in Section 5.

It is noted that the report does not aim to create new scientific knowledge but instead to synthesise relevant studies and insights on open science in a way that can be useful and usable in TRANSIENCE. Similarly, it does not aim to present an extensive literature review to identify every open-science practice available. Instead, D3.8 operationalises major open-science principles such as the FAIR data principles as well as insights from previous research studies that have synthesised open-science practices in the topics of TRANSIENCE. Finally, to further facilitate the use of the guidelines suggested in this report, the information given in the next sections is mainly structured through tables and figures instead of long blocks of text.

---

[3] https://doi.org/10.5281/zenodo.12624920
[4] https://doi.org/10.5281/zenodo.14034437
[5] https://zenodo.org/communities/transience
[6] https://github.com

Funded by the European Union

2 Open science guidelines

The main objective of TRANSIENCE is to explore interactions between decarbonisation and circularity in the context of European industries and create enhanced modelling capacities to support this effort. As such, the project's research activities can be grouped into four categories: 1) conduct participatory and transdisciplinary research to understand stakeholder needs for new modelling capacity (WP2); 2) characterise circularity & decarbonisation options under the lenses of industrial ecology, theories of innovation, and other perspectives (WP3); 3) enhance, open, and interconnect models that can inform the transition to a green and circular industry (WPs 4 & 7); and 4) conduct scenario modelling and case studies in four industrial clusters in Europe based on the enhanced modelling framework and on continuous co-creation with relevant stakeholders (WPs 8 & 11).

Table 1 provides more details on the inputs, processes, and outputs of these research activities and a high-level mapping of these activities with relevant open science principles and guidelines. Specifically, all research activities will be described through open-access publications (deliverable reports, journal papers, etc.) while their input and output data will be shared based on FAIR data principles. Model code as well as relevant scripts (e.g., for data processing and visualisation) will be shared as open source and deposited in online repositories that follow the TRUST principles (Transparency, Responsibility, User focus, Sustainability, and Technology) such as Zenodo.

Table 1. TRANSIENCE research activities and relevant open science principles

<table>
<tr><th colspan="2">TRANSIENCE research activities</th><th>Relevant open science principles</th></tr>
<tr><td rowspan="3">Understanding stakeholder needs for new capacities (WP2)</td><td>Inputs: workshops and interviews with stakeholders</td><td rowspan="8">Document all processes through open-access publications (deliverable reports, scientific publications, media articles)

Share raw and processed data (e.g., model inputs and outputs in WPs 4 and 7 or workshop minutes in WP2) along with databases based on the FAIR data principles

Provide as open source the MIC3 model (simplified and full version), its modules, and all relevant scripts (e.g., pathways? for data processing)</td></tr>
<tr><td>Processes: Three Horizons dialogue framework, cognitive mapping, system dynamics</td></tr>
<tr><td>Outputs: reports, stakeholder-informed research questions</td></tr>
<tr><td rowspan="3">Characterising circularity & decarbonisation options (WP3)</td><td>Inputs: academic and grey literature on circularity, decarbonisation, relevant modelling frameworks, and TRANSIENCE's case study regions</td></tr>
<tr><td>Processes: circularity frameworks, sociotechnical analysis, policy analysis, competitiveness analysis</td></tr>
<tr><td>Outputs: reports, database of circular economy measures</td></tr>
<tr><td rowspan="2">Development of modules & integration into the final model (WPs 4, 7 and Task 12.4)</td><td>Inputs: existing modelling code, stakeholder needs from WP2, academic and grey literature on products and services for circularity and decarbonisation</td></tr>
<tr><td>Processes: development of new modelling capacities, model soft-linking through automated routines</td></tr>
</table>

Funded by the European Union

PROCESSES4PLANET

	Outputs: new open modules, databases, full MIC3 model, simplified modelling tool, new industrial transition scenarios for the EU, national and local levels	Store data and code using **TRUSTworthy data repositories**
Scenario co-creation and analysis; model validation by the stakeholders (WPs 8, 11)	**Inputs**: scenario definitions based on stakeholder engagement activities of the project	
	Processes: model scenarios using the new modules and the integrated MIC3 model	
	Outputs: modelling results in terms of GHG emissions, material needs, costs, industrial energy consumption mix, investment needs at the EU. National and local level (for case studies) etc.	

All principles are presented in Subsections 2.1-2.4 along with examples of implementing them in the context of TRANSIENCE. As there are many studies that explored these principles in detail, the following subsections aim to be concise and distil practical recommendations for TRANSIENCE. It is also acknowledged that there may be overlaps between the principles (e.g., certain FAIR data principles relate to open-access publishing), while there are other principles and practices that could have been explored. However, as mentioned in the Introduction, the selected sets of principles are recognised as influential for open science (Abdelrahman, 2020; daSilvaSantos et al., 2025; Mons et al., 2017) and align well with the research activities of TRANSIENCE, thus forming a good-enough basis for establishing guidelines and protocols for the project.

2.1 FAIR data principles

As the value of open data continues to grow across scientific disciplines, the FAIR Principles emerged as a response from a collective of stakeholders spanning academia, industry, and public institutions. These principles—Findability, Accessibility, Interoperability, and Reusability—were developed to serve as benchmarks for assessing the openness and usability of scientific data (Wilkinson et al., 2016). Rather than dictating exact technical standards, the FAIR framework offers a flexible set of criteria designed to guide the enhancement of data reusability across various implementation contexts. By fostering the reuse of existing data, the principles aim to accelerate scientific progress and innovation. Over time, the FAIR principles have gained traction among major institutions, including the European Commission (Mons et al., 2017). The FAIR principles have been also used in climate-economy research, most recently through the data management activities in the context of the 6th Assessment Report (AR6) of the Intergovernmental Panel for Climate Change (IPCC) (Stockhause et al., 2024). Table 2 presents the main requirements for each FAIR principle along with examples for applying them. These examples have been inspired by the open science protocols developed in the context of the Horizon Europe project IAM COMPACT[7] (Rodés-Bachs et al., 2023, 2024) and have been specifically selected as they convey the gist of each principle while presenting practices that can be also applied in TRANSIENCE.

[7] https://www.iam-compact.eu/

Funded by the European Union

Table 2. FAIR principles and examples of their applications

FAIR principles	Main requirements	Examples of applying the principles
Findability	Data is stored in a way that it is easy for all interested parties to find it.	- Allocate persistent identifiers (PIDs) such as DOIs. - Create searchable, machine-readable metadata (e.g., by submitting datasets in repositories such as Zenodo[8])
Accessibility	Data is readily available by both humans and machines.	- Provide metadata in formats/platforms that are available and free-of-charge (e.g., the ARGOS[9] platform from OpenAIRE), even if the data that they are describing are not open access. - If the data is not fully open, specify explicitly the conditions under which the data are available (Mons et al., 2017) or provide alternatives, e.g., aggregated data when raw data cannot be shared for data protection reasons.
Interoperability	Data is provided in standardised formats that can be read and used by different applications (usually within the scientific domain related to the data).	- Share small datasets in non-proprietary formats such as csv and large ones in open database formats (e.g., PostgreSQL[10]). - Provide standardised ways for data exchange such as application programming interfaces (APIs). - Use metadata vocabularies that follow the FAIR principles, e.g., mentioning at least the author(s), a description, the publication date, the repository, the license, and whether there is an embargo period[11].
Re-usability	Data is formatted and documented in a way that enables effective re-use by different users.	- Provide a license that enables re-usability under well-defined conditions, e.g., CC BY 4.0[12]. - Offer detailed descriptions of the origin and quality of the data, e.g., how it was collected, how it was processed, etc. - Use standardised data structures that ensure transparency and comprehensibility, e.g., based on the Tidy Data format (Wickham, 2014).

Note: The descriptions of the principles are adapted from Wilkinson et al. (2016).

2.2 TRUST principles for digital data repositories

Following the example of the FAIR principles for managing and sharing research data, the TRUST principles are guidelines for the digital repositories that host these data (Lin et al., 2020). TRUST stands for the principles of Transparency, Responsibility, User Focus, Sustainability, and Technology, which are deemed as essential for ensuring reliable and trustworthy data repositories. As shown in Table 3, repositories need to be transparent about their services and responsible for ensuring the authenticity of their data holdings.

[8] https://zenodo.org
[9] https://argos.openaire.eu/splash/
[10] https://www.postgresql.org
[11] https://www.openaire.eu/how-to-comply-with-horizon-europe-mandate-for-rdm
[12] https://creativecommons.org/licenses/by/4.0/deed.en

Repositories also need to cater for their users, including data producers and users from academia, industry, policy, and civil society. Critically, repositories need to have a long-term plan and funding to offer their services and provide the adequate technological infrastructure to enable these services. Table 3 provides some practical examples of applying these principles, such as declaring clear terms of use to ensure transparency. The TRUST principles will be used in two aspects of TRANSIENCE. First, we aim to follow them in the planned upgrades of the IAM PARIS platform (Tasks 8.3 and 11.3)[13], ensuring that the platform can act as a TRUSTworthy repository of modelling data and scenarios for TRANSIENCE and for all other projects that use it. Second, we will publish all data, model-based scenarios, and publications developed in TRANSIENCE in the Zenodo repository, which is expected to fulfil most of the requirements of the TRUST principles.

Table 3. TRUST principles and examples of their applications

TRUST principles	Requirements for repositories	Examples of applying the principles
Transparency	To be transparent about specific repository services and data holdings that are verifiable by publicly accessible evidence.	- Declaring clear terms of use, both for the repository and for the data holdings. - Providing minimum digital preservation timeframe for the data holdings.
Responsibility	To be responsible for ensuring the authenticity and integrity of data holdings and for the reliability and persistence of its service	- Adhering to the designated community's metadata and curation standards - Providing data services e.g. portal and machine interfaces, data download, or server-side processing. - Managing the intellectual property rights of data producers, the protection of sensitive information resources, and the security of the system and its content.
User focus	To ensure that the data management norms and expectations of target user communities are met	- Implementing relevant data metrics and making these available to users. - Providing (or contributing to) community catalogues to facilitate data discovery. - Monitoring and identifying evolving community expectations and responding as required.
Sustainability	To sustain long-term services and preserve data holdings	- Planning sufficiently for risk mitigation, business continuity, disaster recovery, and succession. - Securing funding to enable ongoing usage and to maintain the desirable properties of the data resources. - Providing governance for necessary long-term preservation of data so that data resources remain discoverable, accessible, and usable in the future.

[13] https://iamparis.eu

Funded by the European Union
PROCESSES4PLANET

Technology	To provide infrastructure and capabilities to support secure, persistent, and reliable services.	- Implementing relevant and appropriate standards, tools, and technologies for data management and curation. - Having plans and mechanisms in place to prevent, detect, and respond to cyber or physical security threats.

Note: The principles, requirements, and examples are reproduced from Lin et al. (2020).

2.3 Open-access publishing

Open-access publishing refers to the practice of making scientific publications available to the public under well-defined conditions that permit access, reuse, redistribution, adaptation, and repurposing (UNESCO, 2022). This is usually enabled through appropriate open-access licences such as the ones provided by the Creative Commons organisation[14]. The open-access movement has a long history and has been increasingly popular during the last two decades. There are many successful examples of initiatives in multiple countries that aim to support and finance open-access publications (especially when based in publicly funded research) and transform the publishing model of scientific journals (Demeter et al., 2021). The European Commission has also been part of this effort, providing guidelines for promoting open access to publications and data stemming from EU-funded research (European Commission, 2017). Figure 2 shows the suggested routes for disseminating research in the Horizon 2020 programme, promoting open access but also providing flexibility to the researchers.

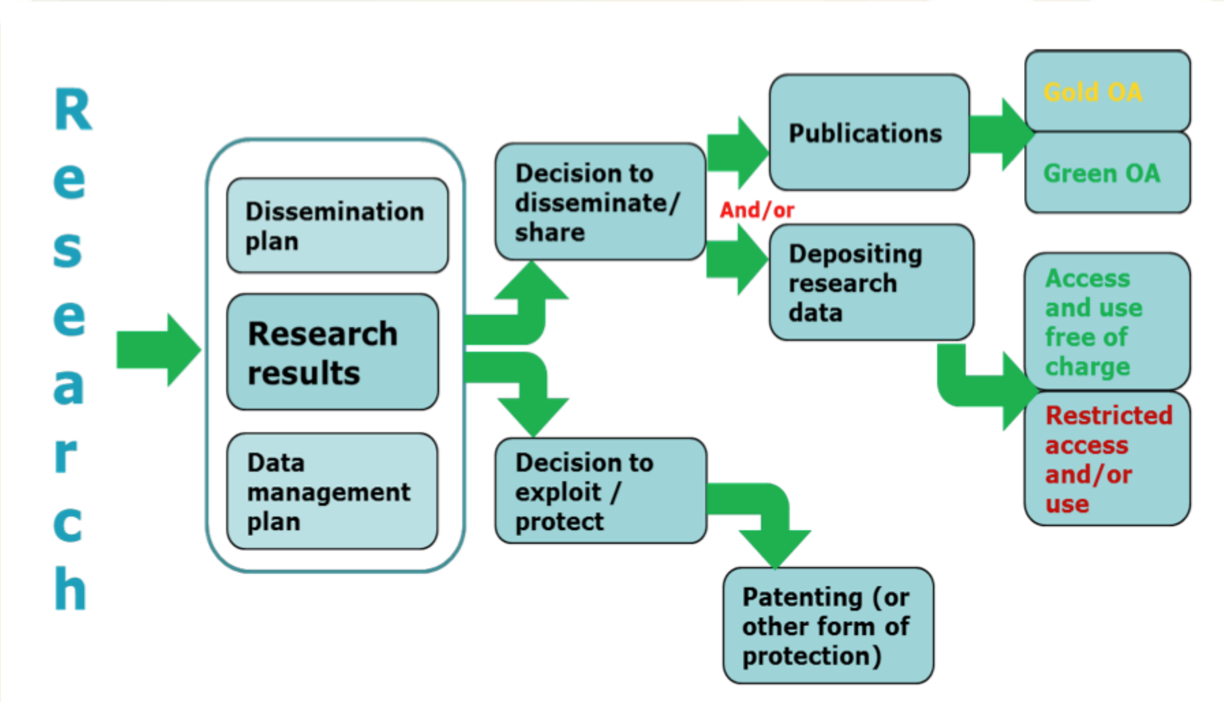

Figure 2. Options for dissemination and exploitation of publications and data in Horizon 2020 projects, reproduced from the European Commission (2017).

[14] https://creativecommons.org

Researchers in Horizon 2020 projects retained the right to exploit research results, e.g., through patenting; however, if they wanted to disseminate their research, they needed to follow the open access rules of the programme[15]. Specifically, the Horizon 2020 programme would cover the costs for publishing scientific articles in Gold Open Access (immediate open access to everyone) or at least Green Open Access (paywall for some months or years), provided that researchers also released the accepted manuscript in an institutional repository. Similar publication routes exist in the Horizon Europe programme (that funds TRANSIENCE) although some of the rules are stricter than Horizon 2020. Most importantly, Horizon Europe will not cover the costs of a publication unless it is published in a fully open-access journal (European Commission, 2025), which is any journal that solely provides the Gold Open Access option and there is no paywall for any of its hosted publications. Nevertheless, researchers can still publish in a hybrid journal that has the Gold Open Access option, provided that the article processing charges (APCs) are paid by the organisation affiliated to the corresponding researcher(s), usually through publishing agreements. A third option is to publish in an open-access journal that does not charges APCs. An example is Open Research Europe, a journal that was recently launched by the European Commission and is specifically designed to host results from EU-funded research projects. All these options are also available for TRANSIENCE's publications and are summarised in Table 4. A similar table has been included in the Data Management Plan of the project (Deliverable D3.1) but without presenting the third option of Open Research Europe that is discussed here.

Table 4. Open access publishing options for TRANSIENCE

Access type	Funded by	Fees	Licenses
Publish in a fully Open Access journal	Horizon Europe Grant	Article Processing Charges; ranging between ~200€ (e.g., Elsevier Societal Impacts[16]) to over 10,000€ (e.g., Nature[17])	CC BY 4.0[18] CC BY-NC-ND 4.0[19]
Publish in a journal that has the option of Gold Open Access	Publishing agreements between the organisations of project partners and the publishers		
Publish in Open Research Europe or a fully open access journal that does not charge APCs[20]	No funding required, Open Research Europe is free for European Commission grant holders.	No fees	

Note: The information for the table has been adapted from OpenAIRE[21].

[15] https://intellectual-property-helpdesk.ec.europa.eu/news-events/news/open-access-obligations-horizon-europe-what-are-cc-licences-2021-11-15_en
[16] https://www.elsevier.com/about/policies-and-standards/pricing
[17] https://www.nature.com/nature/for-authors/publishing-options
[18] https://creativecommons.org/licenses/by/4.0/deed.en
[19] https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en
[20] https://doaj.org
[21] https://www.openaire.eu/how-to-comply-with-horizon-europe-mandate-for-publications

Funded by the European Union

2.4 Open-source software

Closely related to the practices of open access, open source not only suggests that software should be open, but that its source code should be open too. One of the most comprehensive definitions of open source has been compiled by the Open Source Initiative (2025) and presented in Table 5. There are 10 criteria for evaluating whether a software can be assumed as open source, including the provision of its source code and requirements for its license, such as allowing free redistribution and clear terms for modifications and derived works. Table 5 also includes examples of applications that have been selected to match potential use cases in TRANSIENCE, such as creating a data processing script and publishing it as a Python package, allowing it to be redistributed and imported by other programs.

Table 5. Definition of open source from the Open Source Initiative.

Criteria	Description	Examples of application
Free redistribution	The license shall not restrict any party from selling or giving away the software as a component of an aggregate software distribution containing programs from several different sources.	Using a Python package that is shared with the MIT license[22] for developing a model and then sharing the model.
Source code	The program must include source code and must allow distribution in source code as well as in compiled form.	Sharing the source code in a GitHub repository along with a binary release (if applicable)[23].
Derived works	The license must allow modifications and derived works and must allow them to be distributed under the same terms as the license of the original software.	When forking a code repository in GitHub as a starting point for your own software[24]. Some licenses, such as GNU GPLv3, specifically require users to use the same license (strong copyleft)[25] while others, such as the MIT license, are more flexible.
Integrity of the author's source code	The license may restrict source code from being distributed in modified form only if the license allows the distribution of "patch files" with the source code for the purpose of modifying the program at build time.	For instance, the license used for the TeX software[26] stipulates that the original code should not be modified and any further change should be provided in the form of patches.
No discrimination against persons or groups	The license must not discriminate against any person or group of persons.	Avoid restricting the use of a model to specific countries.

[22] https://mit-license.org
[23] https://docs.github.com/en/repositories/releasing-projects-on-github/managing-releases-in-a-repository
[24] https://docs.github.com/en/pull-requests/collaborating-with-pull-requests/working-with-forks/fork-a-repo
[25] https://choosealicense.com/licenses/gpl-3.0/
[26] https://en.wikipedia.org/wiki/TeX

No discrimination against fields of endeavour	The license must not restrict anyone from making use of the program in a specific field of endeavour.	Avoid restricting a model from being used by businesses. Practically, this means that businesses should be allowed to use the model but not sell its use to others. Commercial licenses can be provided based on dual licensing[27].
Distribution of license	The rights attached to the program must apply to all whom the program is redistributed to without the need to execute an additional license by those parties.	Avoid closing the model by any other means, e.g., a non-disclosure agreement.
License must not be specific to a product	The rights attached to the program must not depend on the program's being part of a particular software distribution.	A Python package that is used by a model (and thus distributed by it) retains the same rights as if it is used as a stand-alone program.
License must not restrict other software	The license must not place restrictions on other software that is distributed along with the licensed software	The license cannot insist that programs distributed on the same medium must be open-source software. Copyleft licenses such as GNU GPLv3 stipulate that *derived* software should use the same license and does not apply to software that just distributes the original program.
License must be technology-neutral	No provision of the license may be predicated on any individual technology or style of interface.	Software can be redistributed online or through non-Web channels.

Note: The criteria and descriptions have been reproduced from the Open Source Initiative[28]. The examples are our own elaboration or adapted by the same source.

As with open-access publications, licensing plays an important role in open source too. Nevertheless, the need to share the source code and its derivatives necessitate the use of other licenses than the ones used for open-access publishing, as also recommended by Creative Commons[29]. Figure 3 shows three typical use cases in software development along with respective suggestions for appropriate licenses.

[27] https://en.wikipedia.org/wiki/Multi-licensing
[28] https://opensource.org/definition-annotated
[29] https://creativecommons.org/about/software/

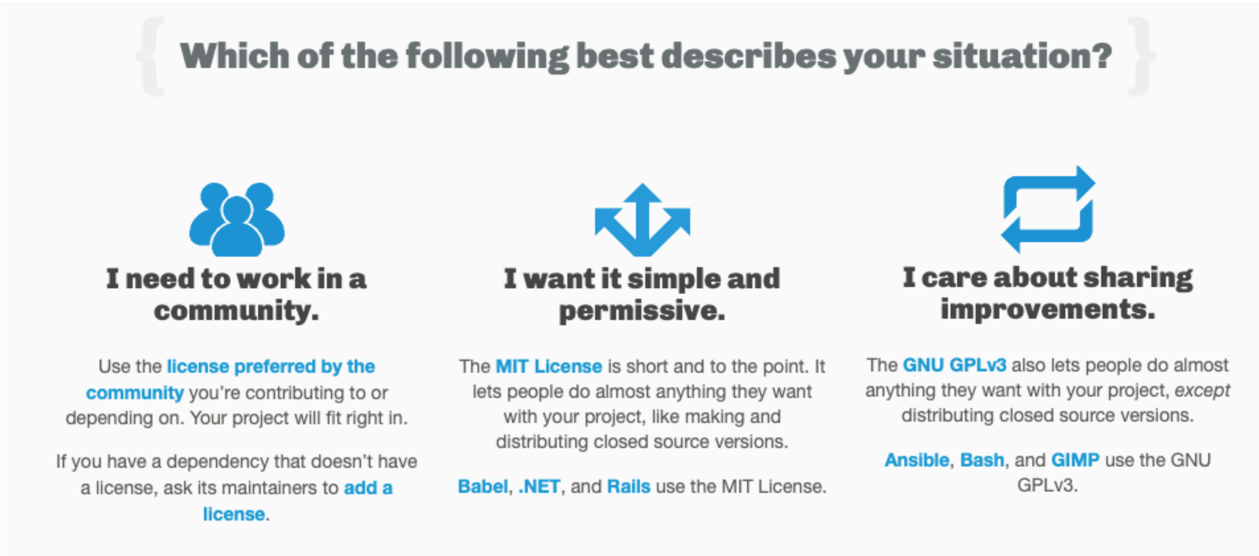

Figure 3. Screenshot from choosealicense.com presenting common use cases when sharing software

While hundreds of licenses exist to match different needs, licenses such as the MIT license[30] and the GNU GPLv3[31] are among the most popular. A variation of the GNU GPLv3 has been used for most of the TRANSIENCE models, as shown in Table 6. The GNU AGPL 3.0[32] is a so-called strong copyleft license as it allows others to modify the code as long as the modifications are released under the same or similar license. The licenses used by the Products & Services (P&S) Database and the Life Cycle Assessment (LCA) model (CC BY 4.0 and BSD 3-Clause, respectively) are much more permissive for re-use, as they only require retaining the original copyright notice or credit the original developers.

Table 6. Open-source licenses and code repositories in TRANSIENCE models

Module	Partner	Focus	License	Code repository
OPEN-GEM	E3M	Socioeconomic impacts	GNU AGPL 3.0[32]	https://github.com/e3modelling/OPEN-GEM
P&S Database	PSI	Products & services	CC BY 4.0[33]	https://zenodo.org/records/15517592
EU MFA	WI & Fraunhofer	Material and sectoral flow in EU	GNU AGPL 3.0[32]	https://github.com/wupperinst/transience-eu-mfa/
ITOM	WI	Industrial	GNU AGPL 3.0[32]	https://github.com/wupperinst/itom

[30] https://choosealicense.com/licenses/mit/
[31] https://choosealicense.com/licenses/gpl-3.0/
[32] https://www.gnu.org/licenses/agpl-3.0.html
[33] https://creativecommons.org/licenses/by/4.0/legalcode

		transformation optimisation		
FORECAST -Sites	Fraunhofer	Industrial technology diffusion	GNU AGPL 3.0[32]	https://github.com/fraunhofer-isi/forecast-sites
REMIND-MFA	PIK	Global material flow and trade	GNU AGPL 3.0[32]	https://github.com/pik-piam/remind-mfa
OPEN-PROM	E3M	Energy system	GNU AGPL 3.0[32]	https://github.com/e3modelling/OPEN-PROM
LCA	PSI	Environmental life cycle assessment	BSD 3-Clause[34]	https://github.com/tomterlouw/lca_transience

Detailed model names: OPEN-GEM = Open General Equilibrium Model for Economy-Energy-Environment, P&S Database = Product & Service Database, EU MFA = Material Flow Analysis for the European Union, ITOM = Industry Transformation Optimization Model, FORECAST-Sites = Technology Diffusion Model in Energy-Intensive Industries, REMIND-MFA = Regional Model of Investment and Development - Material Flow Analysis, OPEN-PROM = Open Prometheus, LCA = Environmental Life Cycle Assessment Model.

[34] https://opensource.org/license/BSD-3-clause

Funded by the European Union

PROCESSES4PLANET

3 Open science protocols for modelling activities

The development, testing, and use of the MIC3 model are among the core objectives of TRANSIENCE. Based on the principles and guidelines presented in Section 2, this section outlines a first version of the protocols that will guide all modelling activities of the project. The role of these protocols is to pinpoint open science practices for each modelling activity and suggest examples and ways that these practices can be applied in the project. For instance, the protocol operationalises the FAIR principles by suggesting modellers to upload model data in Zenodo, provide them with a DOI, and include adequate metadata. In that sense, the protocols of this deliverable are not the strict rules that their name implies, but more of checklists or guidelines, providing an indication to modellers about entry points for open science in their work, while offering them the flexibility to use the method or tool that best fits their model. Nevertheless, these initial protocols will be discussed with all modelling teams during RP2 and adapted to reach a set of minimum practices that will be used for all modelling processes in the project, as a form of quality assurance.

There are many ways to split the modelling process into different activities (Fisher-Vanden & Weyant, 2020; McGookin et al., 2024). Here we divide modelling activities in TRANSIENCE into three main phases: model development, model linking, and model use. These phases roughly correspond to the division of modelling work in TRANSIENCE: during the first reporting period (RP1) modelling teams developed and enhanced the open-source models which will be linked in RP2 to create the MIC3 open modelling framework and, in RP3, used to analyse industrial transition scenarios at EU and Member State levels as well as in selected case studies of industrial complexes. The protocol for each phase is presented in the following sections (Sections 3.1-3.3).

The protocols have been inspired by the protocols created in the IAM COMPACT project (Rodés-Bachs et al., 2023, 2024) and were informed by studies that synthesised and enhanced recommendations from the literature in terms of improving model development (Xexakis & Koasidis, 2025) and model linking (Keppo et al., 2025). As these studies have only been recently completed, they have not yet been fully peer-reviewed by external reviewers. Specifically, the IAM COMPACT protocols are currently in revision after a first round of reviews in the Open Research Europe journal (Rodés-Bachs et al., 2025), Keppo et al. (2025) is under review in the One Earth journal, while a paper based on the study of Xexakis & Koasidis (2025) is under preparation. Nevertheless, all studies have been checked by multiple internal reviewers and co-authors; for instance, the study of Keppo et al. (2025) focusing on model linking includes 20 co-authors with many relevant publications on the topic. Still, we plan to closely follow the revisions in these studies and update the protocols as needed.

3.1 Protocol for model development

This protocol refers to activities that are essential for developing a new model, such as coding, managing data inputs, and writing documentation. Additionally, it includes activities that can directly affect development, such as eliciting modelling requirements from stakeholders and fostering a collaborative environment where the wider modelling community can suggest changes and improvements. The protocol is inspired by the study of Xexakis & Koasidis (2025) on software engineering practices that can be used to improve integrated assessment modelling, including many practices that relate to open science. Additionally, recommendations in terms of model documentation are based on the open science protocols used in the IAM COMPACT project (Rodés-Bachs et al., 2024). The full checklist of the IAM COMPACT protocol and the practices explored in Xexakis & Koasidis (2025) are shown in the Appendix.

Funded by the European Union

PROCESSES4PLANET

Table 7 shows the protocol for model development in TRANSIENCE. Examples of relevant practices include sharing the model code and updates in GitHub, sharing data through the TRANSIENCE community in Zenodo[35], and documenting the model in a Read the Docs website and in the IAM PARIS platform. Some of these suggestions have been already implemented upon the submission of this deliverable report, such as the creation of public GitHub repositories for all modules of MIC3 (see Section 2.4). Other suggestions will be implemented during Phase 2 of TRANSIENCE, such as the creation of Docker images for facilitating model runs.

Table 7. Protocol for ensuring open science in model development

Activity	Open science practices	Suggestions for TRANSIENCE
Elicit modelling requirements	Document the stakeholder elicitation process and its results.	Details of stakeholder workshops and interviews and the resulting research questions documented in an open access publication (already done in D2.2[36]).
Code the model	Share model code.	- Create a dedicated GitHub repository to share each model (already done in RP1 for all modules of MIC3). - Provide links to the code repository in the project website and the IAM PARIS platform[37]. - Share every new model version as a GitHub release[38]. - Link the repository with Zenodo[39] to get DOIs for every release.
Document the model	Document main inputs, outputs, assumptions, equations, and differences between model versions.	- Document the model in the IAM PARIS platform, based on a specific template (see an example of a filled template here[40]). The final template for TRANSIENCE will be developed in the context of Task 7.3. - Provide a summary of the documentation on GitHub. - if preferred, create a detailed Read the Docs website (see an example for the MARIO software[41]).
	Share case studies and tutorials for the model.	Share through publications (deliverables, scientific papers) or through a Read the Docs website.
Manage data	Document (and share)	- Use the data template provided by the Integrated

[35] https://zenodo.org/communities/transience/
[36] https://zenodo.org/records/14751081
[37] https://iamparis.eu/detailed_model_doc/30
[38] https://docs.github.com/en/repositories/releasing-projects-on-github/managing-releases-in-a-repository
[39] https://docs.github.com/en/repositories/archiving-a-github-repository/referencing-and-citing-content
[40] https://i2am-paris.eu/static/excel_files/I2AM%20PARIS_documentation_template_GCAM.xlsx
[41] https://mario-suite.readthedocs.io/en/latest/intro.html

Funded by the European Union

inputs	data inputs.	Assessment Modelling Consortium (IAMC)[42]. - Upload input datasets in Zenodo (and link them to the TRANSIENCE community[43]) and provide some basic metadata (authors, version, release date, etc.). - If input datasets are not open access, find alternative datasets that are open or explore ways to publicly share them, e.g., using pre-built binary files. - Ideally, provide more detailed metadata, e.g., using the metadata format used in the IPCC scenario databases[44] or the metadata templates used in the project's Data Management Plan in ARGOS[45] (managed by ICCS).
	Share data processing scripts.	Store data processing scripts in GitHub, ideally in the same repository as the model.
Foster collaboration	Provide a space for new model users to ask questions and provide suggestions.	Use the Issues functionality in the GitHub repository for each model[46].
Deploy the model	Document dependencies and requirements.	- Document the main system requirements (e.g., CPU, RAM, disk space, operating system) in the README of the model repository on GitHub. - Document dependencies in the same way and, if possible, use a standardised format or an automation software (e.g., using a pyproject.toml file or the Poetry package in Python-based models[47])
	Share a deployable image of the model.	- Create a Docker image that automates code deployment, downloads dependencies, and provisions the right system requirements. - The Docker image can be then shared in DockerHub[48] to allow anyone to easily run their model in their computer or server.

[42] https://pyam-iamc.readthedocs.io/en/stable/#timeseries-types-data-formats
[43] https://zenodo.org/communities/transience/
[44] https://pyam-iamc.readthedocs.io/en/stable/data.html#meta
[45] https://argos.openaire.eu/explore-plans/overview/public/30df1ce9-5077-4f28-bfdc-31902af8d910
[46] https://docs.github.com/en/issues
[47] https://python-poetry.org
[48] https://hub.docker.com

Funded by the European Union
PROCESSES4PLANET

3.2 Protocol for model linking

The protocol divides the model linking process into three major activities: a first preparatory phase where modellers need to reconcile the different assumptions, drivers, and spatial and temporal granularities of the models that need to be linked; a second phase where all input data are harmonised; and a third phase where the linking is implemented through software development and data exchange routines. This division roughly corresponds to the structure of a similar checklist by Keppo et al. (2025) that was used to guide modellers through important considerations in the model linking process. Here we adapt and expand the recommendations of the checklist that relate to open science such as sharing a full mapping of assumptions, drivers, and epistemic foundations of the models. Table 8 presents the full protocol along with practical details on applying it in TRANSIENCE.

Table 8. Protocol for ensuring open science in model linking

Activity	Open science practices	Suggestions for TRANSIENCE
Assess modelling assumptions	Describe interaction mechanisms between models.	Create a figure showing all interactions between the models, e.g., using a Unified Modelling Language (UML) diagram[49]. See Fujimori et al. (2024) for an example.
	Providing a full mapping of assumptions, drivers, and epistemic foundations for the models to be linked.	The mapping can be documented in an Excel file (to be shared through Zenodo) or in a related publication (deliverable report, scientific paper).
	Document any method used to reconcile the spatial and temporal granularity of the models (downscale or upscale).	Document in a related publication as above and, ideally, share the scripts used for data downscaling or upscaling.
Harmonise inputs	Document modifications to the inputs and the rationale behind them as well as any other assumption related to the harmonisation process.	Document in a related publication (as above).
	Share the harmonised inputs.	Share the harmonised datasets in Zenodo (and link to the TRANSIENCE community).
Establish interconnections	Use standardised data templates and automation processes to establish the links between the modules.	Use the IAMC template for all data transfers between the models. Facilitate data processing through community tools such

[49] https://miro.com/diagramming/what-is-a-uml-diagram/

Funded by the European Union

PROCESSES4PLANET

		as pyam[50] or nomenclature[51].
	Document (and share) model versions used in the interconnection.	Create a release in the GitHub repository of each model to specify the model version used in the interconnection.
	Share scripts, APIs, or other tools used for establishing the links.	Create a separate GitHub repository to share all scripts used to establish the interconnection.
	Share a tutorial or a case study to illustrate the linking.	Add the tutorial in the same GitHub repository as with the interconnection scripts and/or share it in the publication describing the interconnection.
	Share a deployable image of the interlinked models (optional).	Create a Docker image encapsulating all models, system requirements, dependencies, and linking scripts. The Docker image can be then shared in DockerHub[52] to allow anyone to run the interlinked models in their computer or server.

3.3 Protocol for model use

The final protocol refers to activities related to the use of model for scenario modelling and other exercises. For instance, it includes advice on how to standardise model outputs, how to visualise them, and how to validate modelling inputs and outputs and share the results of this validation to enhance transparency and improve confidence in the results by the users. The division of modelling activities and the related open science practices are mostly based on the open science protocol of the IAM COMPACT project (Rodés-Bachs et al., 2024). Examples of these practices include standardising data templates using the IAMC data format[53], using an interactive tool for validating modelling data and scenario results based on and further enhancing the vetting process of IPCC AR6, and visualising modelling results in the IAM PARIS platform.

[50] https://pyam-iamc.readthedocs.io/en/stable/
[51] https://nomenclature-iamc.readthedocs.io/en/stable/
[52] https://hub.docker.com
[53] https://pyam-iamc.readthedocs.io/en/stable/#timeseries-types-data-formats

Funded by the European Union
PROCESSES4PLANET

Table 9. Protocol for ensuring open science in model use

Phase	Open science practices	Suggestions for TRANSIENCE
Output Standardisation	Standardise model outputs according to an established template.	Use the IAMC data template[53].
Model validation	Validate the first model period against observed data; alignment checks of key inputs and outputs, such as demand and supply.	Use the model validation methods suggested in Wilson et al. (2021). The results of the validation can be shared in a relevant deliverable report or scientific publication to enhance the transparency of the model runs.
	Validate results based on the vetting process of IPCC AR6.	Use the vetting tool provided by IAM PARIS[54]. As part of WP7, this tool will be further customised to the needs of the project's models (e.g., add more detailed variables for industrial sectors). See an example of such project-specific tool for the IAM COMPACT project[55]. The results of the vetting can also be downloaded and documented along with the rest of validation activities.
Model Output Upload	Storage of raw output and study-specific results in an accessible location with a DOI.	Store both raw and processed outputs in Zenodo.
	Share analysis code.	Create a dedicated GitHub repository for the analysis of the scenario study and link it to the data in Zenodo.
Visualisation Tools	Provide comprehensive figures to illustrate the results.	Share the figure creation code in the GitHub repository for the study (see above).
	Development of user-friendly visualisation tools for easy exploration of outcomes.	Upload results in IAM PARIS platform[56]. A detailed tutorial on how to do this will be prepared within 2025.

[54] https://validation.i2am-paris.eu
[55] https://validation.iam-compact.eu
[56] https://iamparis.eu/results

4 Ontologies for data sharing

4.1 Ontologies in science

Ontologies are used to describe and represent a shared understanding of domain knowledge and consist of concepts and relationships between them. Formal axioms are used to constrain the interpretation and well-formed use of the defined concepts (Hadzic et al., 2009). Terms are used to designate the concepts and the relationships in the universe of discourse (Gruber, 1993). Essentially, every scientific field has its own ontology.

Researchers may need to design a common ontology for various reasons (Gruber, 1993):

- common understanding of the domain knowledge (knowledge within a community of people committed to a common goal);
- sharing domain knowledge;
- reusing domain knowledge;
- analysing domain knowledge;
- separation of domain knowledge from operational knowledge;
- making domain assumptions explicit.

In a scientific project context, in order to communicate effectively, researchers, stakeholders, software applications, and other agents (such as automated systems that act independently), need to have a common and shared understanding of the terms used in communication. This requires using the same ontology, so that all parties interpret the information in the same way.

Ontologies reflect how a community understands a particular part of the world. All members of a community (such as a scientific project community) must agree on the meaning and representation of the knowledge that is included in the ontology. Definitions can only be added only if there is consensus. No individual can change the ontology based on their personal preference. When a community commits to an ontology, they can use it to share, reuse, and apply knowledge more effectively (Hadzic et al., 2009).

Ontological commitment refers to the agreement on the meaning of the concepts and their relationships within an ontology (Gruber, 1993). When someone or something (a researcher, system, or agent) commits to an ontology, they agree to use the shared vocabulary in a coherent and consistent manner. This includes accepting the rules, constraints, and definitions that are part of the ontology.

Communities can also adopt ontologies developed by others. Several communities may agree on a general ontology with broader terms and fewer constraints, and then develop more specific ontologies based on it (Hadzic et al., 2009). An example of different ontologies between scientific domains is shown in Figure 4.

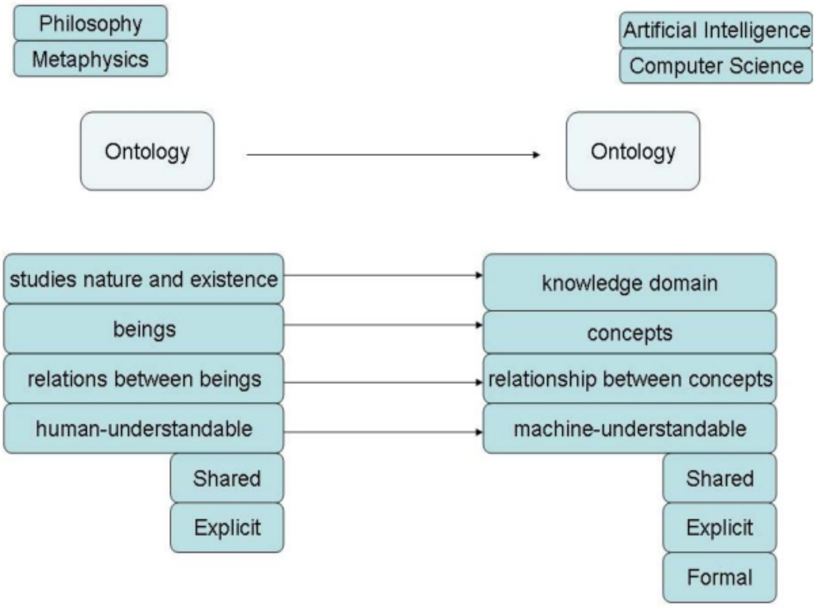

Figure 4. Meaning of ontology in two different contexts, reproduced from Hadzic et al. (2009)

The basic properties of an ontology in the context of computer science and data interoperability are the following (Hadzic et al., 2009):

- They describe a specific domain.
- Ontology users agree to use the ontology terms consistently.
- Ontology concepts and relations are unambiguously defined in a formal language using axioms and definitions.
- Relationships between ontology concepts determine the ontology structure e.g. hierarchical or non-hierarchical. The generalisation/specialisation relationship (i.e. “is-a” relationship) between two concepts is an example of a hierarchical relationship between concepts.
- They can be understood by computers.

Based on these properties, two important functions of ontologies are the following:

1. They enable project ecosystem’s agents to work cooperatively and communicate with each other.
2. They make the available information more accessible to automated agents (Hadzic et al., 2009).

When different systems are based on the same ontology, they can extract, analyse, and use information from each other more easily. Describing web content in a machine-understandable way makes online information more accessible to automated tools (Horrocks, 2002). An ontology usually has a hierarchical structure and defines domain concepts and their properties. It is a collection of shared and precisely defined terms that can be used to describe content in a way that computers can understand. This makes information easier to find, access, and use effectively (Hadzic et al., 2009).

4.2 Ontologies in TRANSIENCE

In the TRANSIENCE project, we will create and share ontologies to document the key elements of the research process, including variables, parameters, concepts, and assumptions used in the modelling activities of the project. The goal is to build a common and structured vocabulary that can be used by all members of the project and by the wider research community. As part of our commitment to open science, these ontologies will enable researchers within the TRANSIENCE project, as well as external stakeholders and the global research and modelling community, to share, analyse, evaluate, and apply the project's produced knowledge in a consistent and transparent way. A first discussion on ontologies has already taken place in the RP1 of TRANSIENCE. The remainder of this section—as well as Section 4.3—summarises some initial points that were discussed.

These ontologies are being designed to cover the basic concepts used in the project, with the goal to support knowledge sharing, improve transparency, and enable interoperability between different models, datasets, and research teams. The first phase of creating the project's ontology involves identifying the key concepts and relationships that arise in the project's domain. This includes the variables and parameters used in the models, the assumptions made in scenario design and building, and the key terms that describe these processes and their results. Where needed, the project will build on existing ontologies and standards to ensure alignment with the broader research community and to ensure interoperability. A fundamental example is the use of IAMC standards for variable and parameter nomenclature by the modelling teams.

The ontologies will be implemented in formal languages like the Web Ontology Language (OWL)[57] or the Resource Description Framework Schema (RDF)[58] so they can be processed by machines. The implementation can be facilitated by appropriate tools like Python libraries rdflib[59] and owlready2[60] as well as Protégé[61] for a more visual approach. The concepts, relationships, and rules will be clearly defined through formal axioms to ensure their consistent use. The final ontologies will be published with full documentation, and made openly available to the wider community, following open science principles. This will support their reuse, adaptation, and further development by other researchers and stakeholders. In this way, TRANSIENCE will enhance its transparency and the reproducibility of its research outputs and will contribute to a more open and interoperable knowledge base in the project's domain.

4.3 Example of ontology in TRANSIENCE's modelling data

We provide an indicative example of transforming TRANSIENCE domain knowledge into a machine-readable ontology structure. Table 10 showcases an example (sub-set) of the variables used in the TRANSIENCE models based on the IAMC template. The information is provided from the project's modelling teams and, although still under heavy development, it enables the creation of a basic ontology structure in OWL. Initially, an Excel file with a table similar to Table 10 is parsed into a YAML file that can be used for variable and parameter assessments such as the ones enabled through IAMC's nomenclature package[62]. We then transform the YAML file into an OWL file (see example in Figure 5), that will be ready to support semantic reasoning and rich relationships. The processing script and the current datasets in Excel, YAML, and OWL

[57] https://www.w3.org/OWL/
[58] https://www.w3.org/2007/OWL/wiki/RDF-Based_Semantics
[59] https://rdflib.readthedocs.io/en/stable/index.html
[60] https://owlready2.readthedocs.io/en/latest/
[61] https://protege.stanford.edu
[62] https://nomenclature-iamc.readthedocs.io/en/stable/

Funded by the European Union
PROCESSES4PLANET

formats can be found in a dedicated repository on GitHub[63].

Table 10. Example of variable and parameter data from TRANSIENCE models

Type	Name	Unit	Definition
parameter_input	Material intensity|[material]|[car type]	kg/car	Material intensity passenger cars
	Stock|Buildings|[building types]|[age cohort]	m2	Building stock development
	Stock|Cars|[car types]	car	Passenger car inflow
parameter_output	Production|Clinker	t/yr	Clinker production
	Production|Plastics|[Product categories]	t/yr	Plastics production
	Production|Steel|[Product categories]	t/yr	Steel production
variable	Capacity|Electricity	GW	Total installed (available) capacity of all operating power plants
	Capacity|Electricity|Hydro	GW	Total installed (available) capacity of hydropower plants.
	Capacity|Electricity|Wind	GW	Total installed (available) capacity of wind power plants (onshore + offshore).
	Capacity|Electricity|Solar|PV	GW	Total installed (available) capacity of solar PV power plants.

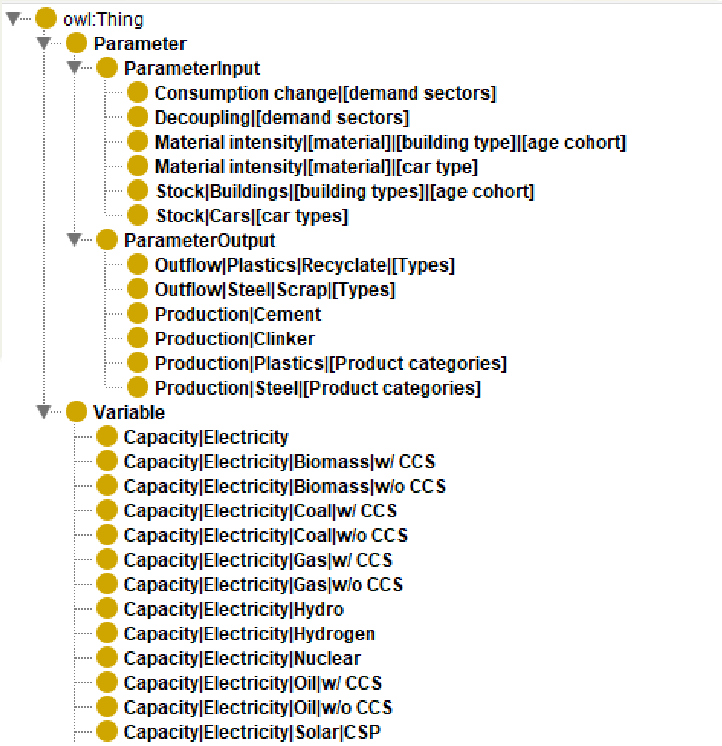

Figure 5. The OWL file tree of the TRANSIENCE project's nomenclature example visualised using the Protégé software.

[63] https://github.com/i2amparis/transience_open_science

Funded by the European Union
PROCESSES4PLANET

5 Next steps

Building on the TRANSIENCE Data Management Plan, this report explores open science principles beyond data sharing and provides practical examples to support their implementation in the project. As emphasised in the introduction, this report does not aim to present an extensive literature review of open science practices but provide practical, hands-on guidelines for the project consortium to follow in their research activities. Since modelling activities form the core of the project, this report provides detailed protocols for model development, linking, and use as well as an application of ontologies to support data sharing. While the first phase of TRANSIENCE included the first set of upgrades of each individual module of TRANSIENCE, the protocols and ontologies will be mostly useful in the second and third phase of the project where the upgraded individual models will be finalised, tested, adapted, and linked together to form the MIC3 model, which will also be, subsequently, tested in different case studies and scenario design assumptions. These phases will also include most of the publications and data that will be produced by the project, making open science practices especially relevant.

The open science protocols and ontologies will be further updated later in the project based on the feedback and experience of the modelling partners in the context of Tasks 7.1, 7.3, and 11.1 and documented in the respective deliverables (D7.1, D7.2, and D11.2). As model linking will be critical for creating the interconnected MIC3 model, the protocols will be enriched with more detailed information about the development of application programming interfaces (APIs) to facilitate these links and the packaging of models and their system requirements in Docker images. The ontology will also be further expanded to include all input and output data from the individual models, facilitating the mapping of the different data transfers that need to take place in MIC3. Additionally, we will aim to standardise and disseminate the TRANSIENCE protocols and ontologies as much as possible to make them useful and usable by modellers beyond TRANSIENCE, e.g., by publishing them in scientific publications and creating spreadsheet-like checklists. Finally, in consultation with the project partners of WPs 8 and 11, we may create additional protocols to cover data sharing related to stakeholder activities; for instance, how to share workshop minutes, transcripts of interviews, or any other raw data from stakeholder interactions that goes beyond reporting results in scientific publications.

Funded by
the European Union

6 References

Abdelrahman, O. H. (2020). Overview of the Principles and Practices of Open Access Publishing. In *Digital Libraries—Advancing Open Science*. IntechOpen. https://doi.org/10.5772/intechopen.95355

daSilvaSantos, L., Burger, K., Kaliyaperumal, R., & Wilkinson, M. (2025). FAIR Data Point: A FAIR-Oriented Approach for Metadata Publication. *Data Intelligence*, *5*(1), 163–183. https://doi.org/10.1162/dint_a_00160

Demeter, M., Jele, A., & Major, Z. B. (2021). The International Development of Open Access Publishing: A Comparative Empirical Analysis Over Seven World Regions and Nine Academic Disciplines. *Publishing Research Quarterly*, *37*(3), 364–383. https://doi.org/10.1007/s12109-021-09814-9

European Commission. (2017). *Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020*. https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

European Commission. (2025). *Open access*. https://research-and-innovation.ec.europa.eu/strategy/strategy-research-and-innovation/our-digital-future/open-science/open-access_en

Fisher-Vanden, K., & Weyant, J. (2020). The Evolution of Integrated Assessment: Developing the Next Generation of Use-Inspired Integrated Assessment Tools. *Annual Review of Resource Economics*, *12*(Volume 12, 2020), 471–487. https://doi.org/10.1146/annurev-resource-110119-030314

Fujimori, S., Oshiro, K., Nishiura, O., Hasegawa, T., & Shiraki, H. (2024). Integration of a computable general equilibrium model with an energy system model: Application of the AIM global model. *Environmental Modelling and Software*, *178*. Scopus. https://doi.org/10.1016/j.envsoft.2024.106087

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, *5*(2), 199–220. https://doi.org/10.1006/knac.1993.1008

Hadzic, M., Wongthongtham, P., Dillon, T., & Chang, E. (2009). Introduction to Ontology. In M. Hadzic, P. Wongthongtham, T. Dillon, & E. Chang (Eds.), *Ontology-Based Multi-Agent Systems* (pp. 37–60). Springer. https://doi.org/10.1007/978-3-642-01904-3_3

Horrocks, I. (2002). DAML+OIL: A Description Logic for the Semantic Web. *IEEE Data Eng. Bull.*, *25*(1), 4–9.

Keppo, I., Khourdajie, A. A., Gardumi, F., Holtz, G., Nikas, A., Xexakis, G., Ferreras-Alonso, N., Fragkos, P., Frilingou, N., Ghadaksaz, H., Hawkes, A., Mateo, A., Mittal, S., Parrado-Hernando, G., Peters, G. P., Rinaldi, L., Rocco, M. V., Sognnæs, I., Ven, D.-J. van de, & Zamanipour, B. (2025). Model linking for low-carbon transitions: Technical and conceptual challenges and best practices. *One Earth*, *under review*.

Lin, D., Crabtree, J., Dillo, I., Downs, R. R., Edmunds, R., Giaretta, D., De Giusti, M., L'Hours, H., Hugo, W., Jenkyns, R., Khodiyar, V., Martone, M. E., Mokrane, M., Navale, V., Petters, J., Sierman, B., Sokolova, D. V., Stockhause, M., & Westbrook, J. (2020). The TRUST Principles for digital repositories. *Scientific Data*, *7*(1), Article 1. https://doi.org/10.1038/s41597-020-0486-7

McGookin, C., Süsser, D., Xexakis, G., Trutnevyte, E., McDowall, W., Nikas, A., Koasidis, K., Few, S., Andersen,

P. D., Demski, C., Fortes, P., Simoes, S. G., Bishop, C., Rogan, F., & Ó Gallachóir, B. (2024). Advancing participatory energy systems modelling. *Energy Strategy Reviews*, *52*, 101319. https://doi.org/10.1016/j.esr.2024.101319

Mons, B., Neylon, C., Velterop, J., Dumontier, M., Da Silva Santos, L. O. B., & Wilkinson, M. D. (2017). Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud. *Information Services and Use*, *37*(1), 49–56. https://doi.org/10.3233/ISU-170824

Open Source Initiative. (2025). *The Open Source Definition*. https://opensource.org/osd

Rodés-Bachs, C., Horowitz, R., & Van de Ven, D.-J. (2024). *IAM_COMPACT_D3.7_Open_Science_Protocols_update*. IAM COMPACT. https://zenodo.org/records/13957129

Rodés-Bachs, C., Sampedro, J., Frilingou, N., Gardumi, F., & Lo Giudice, C. (2025). Open Science Practices in Integrated Assessment Models. *Open Research Europe*, *5*(12). https://doi.org/10.12688/openreseurope.18824.1

Rodés-Bachs, C., Sampedro, J., Horowitz, R., & Van de Ven, D.-J. (2023). *IAM_COMPACT_D3.6_Open_Science_Protocols*. IAM COMPACT. https://zenodo.org/records/10464845

Stockhause, M., Huard, D., Khourdajie, A. A., Gutiérrez, J. M., Kawamiya, M., Klutse, N. A. B., Krey, V., Milward, D., Okem, A. E., Pirani, A., Sitz, L. E., Solman, S. A., Spinuso, A., & Xing, X. (2024). Implementing FAIR data principles in the IPCC seventh assessment cycle: Lessons learned and future prospects. *PLOS Climate*, *3*(12), e0000533. https://doi.org/10.1371/journal.pclm.0000533

UNESCO. (2022). *Online Open Access programme and meeting document Understanding open science*. UNESCO. https://doi.org/10.54677/UTCD9302

Wickham, H. (2014). Tidy Data. *Journal of Statistical Software*, *59*, 1–23. https://doi.org/10.18637/jss.v059.i10

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). Comment: The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*, 1–9. https://doi.org/10.1038/sdata.2016.18

Wilson, C., Guivarch, C., Kriegler, E., van Ruijven, B., van Vuuren, D. P., Krey, V., Schwanitz, V. J., & Thompson, E. L. (2021). Evaluating process-based integrated assessment models of climate change mitigation. *Climatic Change*, *166*(1–2). https://doi.org/10.1007/s10584-021-03099-9

Xexakis, G., & Koasidis, K. (2025). *D6.6 Open-source software development pipelines*. DIAMOND. https://zenodo.org/records/15592781

7 Appendix

7.1 Software engineering practices for modelling development

Table 11. Software engineering practices, reproduced from Xexakis & Koasidis (2025)

Model development	Documentation	Evaluation	Versioning and collaboration	Deployment
Requirements Engineering	Conceptual documentation	Code testing	Version control	Dependency management
Practices for data processing	Code documentation	Model validation	Issue tracking	System requirements management
Programming paradigms, patterns, and standards	Data documentation	Continuous integration	Open-source development	Continuous deployment
Development lifecycle models	Tutorials and case studies			
Practices for refactoring and adding new features				

7.2 Open science protocol used in the IAM COMPACT project

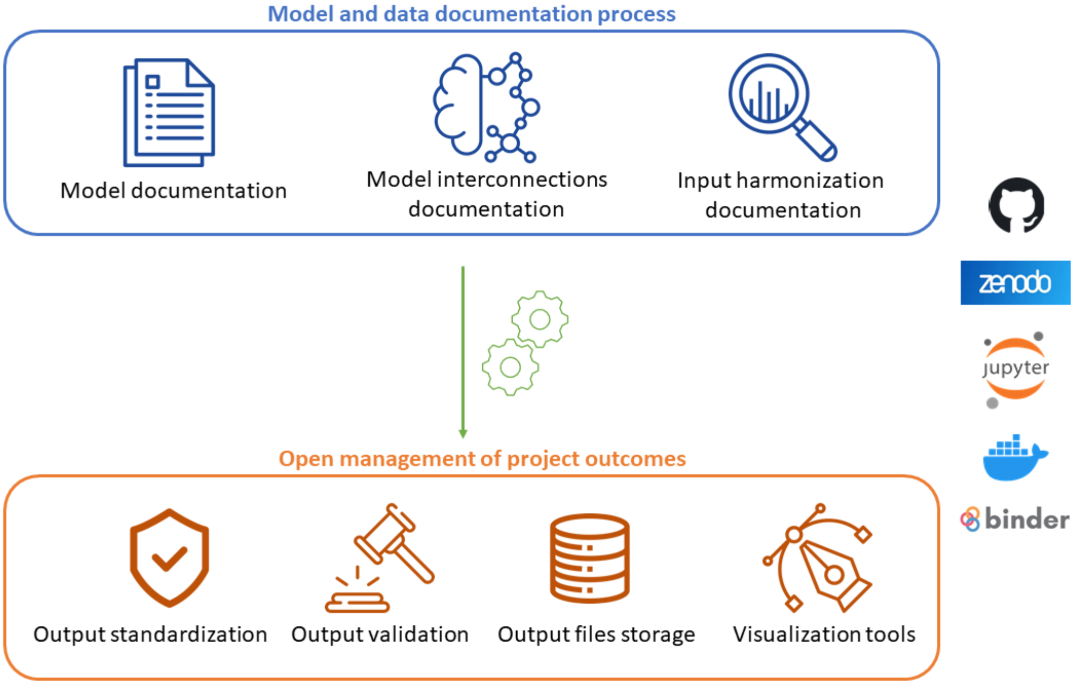

Figure 6. Open science protocol used in IAM COMPACT, reproduced from Rodés-Bachs et al. (2024)

Protocol checklist, reproduced from Rodés-Bachs et al. (2024)

1) **Model and Data Harmonisation Process.**
   a) Model Documentation.
      - A dedicated platform for model documentation, including:
         - Details of inputs.
         - Details of outputs.
         - Details of assumptions.
         - Details of equations.
         - Details of model versions.
      - Assignment of a DOI to model release versions.
   b) Model Interconnections (if applicable)
      - A dedicated section in the paper detailing:
         - Interaction mechanisms between models.
         - Inputs and outputs of the interconnections.
         - Assumptions specific to the interconnections.

        - Model versions used.

   c) Input Harmonisation (if applicable).
      - A dedicated section in the paper detailing:
        - Specific modifications to the inputs and the rationale behind them.
        - Assumptions related to the harmonisation process.
        - Model versions used.
      - Pointer or citation to the specific modified model version, if applicable.

2) Open Management of Project Outcomes.

   a) Output Standardisation.
      - Standardisation of model outputs according to the time-series data template.

   b) Vetting Procedure.
      - Validation of the first model period against observed data.
      - Alignment checks of key outputs, such as demand and supply.

   c) Model Output Upload.
      - Storage of raw output in an accessible location.
      - Storage of study-specific results in an accessible location.
      - Availability of analysis code and figure creation code.
      - Availability of metadata detailing data origin and handling process.
      - Assignment of a DOI to the model outputs and analysis code.

   d) Visualisation Tools.
      - Provision of comprehensive figures to illustrate the results.
      - Development of user-friendly visualisation tools for easy exploration of outcomes.