IAM COMPACT

Expanding Integrated Assessment Modelling: Comprehensive and Comprehensible Science for Sustainable, Co-Created Climate Action

D3.2 - Open data management plan

WP3 – Exchanging – Open & FAIR science, mutual learning

Funded by the European Union

26/12/2023

www.iam-compact.eu

Disclaimer

Copyright Message

Grant Agreement Number	101056306	**Acronym**	IAM COMPACT
Full Title	Expanding Integrated Assessment Modelling: Comprehensive and Comprehensible Science for Sustainable, Co-Created Climate Action		
Topic	HORIZON-CL5-2021-D1-01-04		
Funding scheme	HORIZON EUROPE, RIA – Research and Innovation Action		
Start Date	September 2022	**Duration**	36 Months
Project URL	https://www.iam-compact.eu		
EU Project Advisor	Andreas Palialexis		
Project Coordinator	National Technical University of Athens – NTUA		
Deliverable	D3.2 – Open data management plan		
Work Package	WP3 – Exchanging – Open & FAIR science, mutual learning		
Date of Delivery	**Contractual** 31/12/2023	**Actual**	26/12/2023
Nature	Report	**Dissemination Level**	Public
Lead Beneficiary	Imperial College London (Imperial)		
Responsible Author	Shivika Mittal	**Email**	s.mittal@imperial.ac.uk
	Imperial	**Phone**	-
Contributors	Alexandros Nikas, Natasha Frilingou, Konstantinos Koasidis [NTUA]		
Reviewer(s)	Lorenzo Rinaldi [POLMI], Nikos Kleanthis [UNIPI], Natasha Frilingou [NTUA]		
Keywords	Data Management; Open source; Open access; FAIR data		

EC Summary Requirements

1. Changes with respect to the DoA

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

2. Dissemination and uptake

This report shall serve as a guide for all consortium partners on how to manage datasets and how to update the machine-actionable Data Management Plan (maDMP) in ARGOS, with adequate descriptions and metadata for these datasets. It can be used by entities outside the consortium to explain how data in the IAM COMPACT project is managed, stored, shared, and disseminated.

3. Short summary of results (<250 words)

This report documents the first version of the Open Data Management Plan (DMP) of IAM COMPACT. The DMP is a dynamic framework that will be maintained and modified throughout the project. It currently provides information on data description, a data sharing methodology and resource allocation to achieve Findable, Accessible, Interoperable, and Reusable (FAIR) data, and details on how the project will ensure data security and adherence to ethical standards. Besides this report, IAM COMPACT will be using the ARGOS service of OpenAIRE and EUDAT to deliver a machine-actionable data management plan (maDMP). The maDMP will be continuously updated with metadata for project datasets that will be generated from project activities. The DMP report will also be updated at the end of the project (D3.3).

4. Evidence of accomplishment

This report, as well as the machine-actionable version of the DMP alongside its datasets in ARGOS (link).

Funded by
the European Union

Preface

IAM COMPACT supports the assessment of global climate goals, progress, and feasibility space, and the design of the next round of Nationally Determined Contributions (NDCs) and policy planning beyond 2030 for major emitters and non-high-income countries. It uses a diverse ensemble of models, tools, and insights from social and political sciences and operations research, integrating bodies of knowledge to co-create the research process and enhance transparency, robustness, and policy relevance. It explores the role of structural changes in major emitting sectors and of political, behaviour, and social aspects in mitigation, quantifies factors promoting or hindering climate neutrality, and accounts for extreme scenarios, to deliver a range of global and national pathways that are environmentally effective, viable, feasible, and desirable. In doing so, it fully accounts for COVID-19 impacts and recovery strategies and aligns climate action with broader sustainability goals, while developing technical capacity and promoting ownership in non-high-income countries.

NTUA – National Technical University of Athens	EL	
Aalto – Aalto Korkeakoulusaatio SR	FI	
AAU – Aalborg Universitet	DK	
BC3 – Asociacion BC3 Basque Centre for Climate Change – Klima Aldaketa Ikergai	ES	
Bruegel – Bruegel AISBL	BE	
CARTIF – Fundacion CARTIF	ES	
CICERO – Cicero Senter for Klimaforskning Stiftelse	NO	
E3M – E3-Modelling AE	EL	
KTH – Kungliga Tekniska Hoegskolan	SE	
POLIMI – Politecnico di Milano	IT	
UPRC – University of Piraeus Research Center	EL	
UVa – Universidad De Valladolid	ES	
WI – Wuppertal Institut fur Klima, Umwelt, Energie GGMBH	DE	
IIMA – Indian Institute of Management	IN	
THU – Tsinghua University	CN	
USMF – University System of Maryland	US	
AAiT – Addis Ababa University	ET	
KEI – International Civic Organisation Kyiv Economics Institute	UA	
RUSL – Raja Rata University of Sri Lanka	LK	
TUM – Technical University of Mombasa	KE	
UNIGE – Université de Genève	CH	
Imperial – Imperial College of Science, Technology and Medicine	UK	

Funded by
the European Union

Executive Summary

This report documents the first version of the Open Data Management Plan (DMP) of IAM COMPACT. The DMP describes the data management life cycle for the data to be collected, processed, and generated, and as such it is a dynamic framework to be maintained and modified throughout the project's lifetime. It currently provides information on data description, a data sharing methodology and resource allocation to achieve Findable, Accessible, Interoperable, and Reusable (FAIR) data, and details on how the project will ensure data security and adherence to ethical standards. Besides this report, a key consideration for the data management plan is to render it directly machine-actionable (ma-DMP), enabling automation of data management actions, leading to FAIR data and outputs. To this end, IAM COMPACT will be using the ARGOS service of OpenAIRE and EUDAT to deliver a maDMP. The maDMP will be continuously updated with metadata for project datasets that will be generated from project activities. The DMP report will also be updated in M36 (D3.3).

Funded by
the European Union

Contents

Table of Tables

1 Introduction

This report is the initial version of the Horizon Europe IAM COMPACT project's open Data Management Plan (DMP). The DMP provides guidelines and procedures for collecting, generating, processing, and sharing data throughout the project's duration. Currently, the report includes a description of the expected data and methodologies that will be used for data processing, curation, and preservation (Section 2). The report will cover the entire lifecycle of project data, encompassing data handling during and after the project, along with a comprehensive strategy to ensure that this data is Findable, Accessible, Interoperable, and Reusable (FAIR; see Section 3). This is crucial to adhere to the project objectives of ensuring all project data and outputs (i.e., developed models and their results, publications, as well as the training material utilised for capacity building) are open source (Section 4). In addition to data openness, the report will address security (Section 5) and ethical considerations (Section 6) related to data collection and usage. The DMP will be updated in M36 to document any changes to the project's data handling and processing practices.

In parallel with this document, we have developed a machine-actionable Data Management Plan (maDMP) using the ARGOS[1] service of the OpenAIRE[2]. ARGOS is an online tool in support of automated processes to creating, managing, sharing, and linking DMPs with research outputs. It has been developed by OpenAIRE and EUDAT to deliver an open platform for Data Management Planning that addresses FAIR and Open best practises. It applies common standards for machine-actionable DMPs as defined by the global research data community of RDA and by communicating and consulting with researchers, research communities and funders to better reflect on their needs. The maDMP is an online collection of rich metadata for all datasets that will be developed during the project as well as specific guidelines for each dataset, for instance, on how the dataset will be shared and whether it is linked to another project output, e.g., publications. The datasets themselves will be stored in Zenodo which is a digital repository that follows the desired principles of Transparency, Responsibility, User focus, Sustainability and Technology (TRUST) [1], while the maDMP will provide links to them. The IAM COMPACT maDMP has been created (Version 0) but cannot be finalised due to a bug in the ARGOS platform, for which we have opened a ticket (Ticket ID: 843886, Date: 09/Dec/2023). As soon as the bug is resolved, the maDMP can be finalised and transferred onto IAM COMPACT's Zenodo[3].

Through the ARGOS service, the maDMP can be read by both humans and machines and connect to the OpenAIRE and European Open Science Cloud (EOSC) services, improving the reusability and findability potential of the data. The maDMP will be a "living document" which will be updated immediately after the publication of a project dataset in Zenodo, without having to wait for the annual updates of this report. Nevertheless, these updated reports will provide summaries of all datasets that are currently present in the maDMP. The final chapter of this report (Section 7) provides a first strategy on how new datasets will be added in the maDMP as well as on the roles of the different consortium partners in this process. Based on the experience of the partners with this strategy, adaptations will be made as needed.

[1] https://argos.openaire.eu/home

[2] https://argos.openaire.eu/plans/overview/6510151d-9ab5-46d6-8468-388544c13b42

[3] https://zenodo.org/communities/iam-compact?q=&l=list&p=1&s=10&sort=newest

Funded by the European Union

2 Data Description

2.1 Project objectives and implications for data collection and generation

One of the main objectives of IAM COMPACT is to evaluate global progress towards achieving climate neutrality by the mid-century, by assessing the national policies and needs for the post-2030 world. We will use diverse IAM and sectoral models to assess the regional and national scenarios. We plan to use different IAM and sectoral models to evaluate regional and national scenarios. A crucial aspect of IAM COMPACT is to involve stakeholders in the process of policy-relevant modelling by developing scenarios based on their priorities in the climate debate The project seeks to achieve policy relevance by engaging stakeholders at each step of the process, from framing research questions to building scenarios, and discussing and disseminating results.

Apart from this, we will also co-create and develop scientific and technical capabilities in selected countries and support them in designing their national pledges. As part of this capacity building process, several sectoral and energy system models will be developed for the selected countries, including Sri Lanka, Ethiopia, Kenya, and Ukraine. The project aims to achieve specific objectives through data collection and generation that will involve developing training materials for capacity building, creating new models, and utilising existing IAM and sectoral models to address policy-related questions through stakeholder interactions. To ensure transparency, all new model code and training material (documented in MS6 – "Training material on concepts & tools") will be open source. In case of multi-model assessments, we will compile the data for the harmonisation of key inputs among the various models (details about the harmonisation process and data are mentioned in Deliverable 4.3 - "Broad Scenario Logic"). Based on the priorities and needs of project stakeholders, we will develop scenarios to be run by the models, generating new modelling results. These results will be communicated to all interested parties through scientific publications, policy briefs, and business guides. The research questions and feedback of the stakeholders are documented to ensure the transparency (details are mentioned in Deliverable 2.2 – "Scoping policy relevant research questions").

All outputs of scenario analysis and model documentation will be hosted in the data sharing platform I²AM PARIS[4]. The platform was developed by project coordinator NTUA, project partners BC3, and HOLISTIC (outside the IAM COMPACT consortium) in the context of the H2020 PARIS REINFORCE project, with the aim to establish a portal for the modelling community in support of climate policy and research. Currently, the platform allows modelling teams to host detailed documentation for their models and to develop interactive visualisations to describe the results of modelling exercises.

[4] https://i2am-paris.eu

Funded by the European Union

2.2 Types and formats of data to be generated and collected

Each of the project activities described in Section 2.1 will require and generate different kinds of data. Table 1 shows a non-exhaustive list of the data types and formats that will be used per project activity. Data used for modelling inputs will be mostly quantitative and based on datasets in spreadsheet format (e.g., xlsx or csv) or prebuilt data (e.g., Rda format used for the R language). Regarding modelling development, new model code will be in the format of the programming languages used to develop these models. For example, MicrogridsPy development will primarily be based on the Python language, and its code will be documented in the respective file formats. Qualitative insights from previous reports and publications and from stakeholder co-creation activities will be also used to inform modelling. As with most data inputs, the outputs of the scenario analysis performed in IAM COMPACT will also be quantitative and exported in a spreadsheet format or other formats that are easily converted to spreadsheets. These output datasets will inform the scientific and policy publications of the project which will be provided in pdf format.

Table 1. Data types and formats per project activity

Project activity	Data collected and generated	Data formats
Stakeholder engagement as part of policy response mechanism		Reports documenting stakeholder feedback in pdf
Model documentation and development of new models	**Collect:** Documentation of model characteristics **Generate**: Modelling seminar videos and previous model documentation and new models	html, xlsx, mp4, pdf
Compile input data for the models for the harmonisation	**Collect**: Socio-economic data, historical emissions data, technological cost data, energy use data, capacity data, policy database, National target data and other modelling inputs	xlsx, csv, txt, json and process code is available on github https://cicerooslo.github.io/iam_compact_D43_SSP2_compare/
	Generate: Curate the raw data into a format and data structure that is usable by project models	
Scenario analysis using the stakeholder feedback	**Collect**: During the policy response mechanism, research questions and scenario narratives are developed based on stakeholder engagement.	Report documenting the feedback from stakeholders
	Generate: GHG emissions, jobs created, costs and other social, economic, and environmental indicators	xlsx, csv, txt, json

2.2.1 Data processing tools

Models will be the main data processing tools used in the project. Based on their data inputs, the new models along with existing and established IAM and sectoral models will process the scenarios and narratives suggested by IAM COMPACT stakeholders to provide quantitative results for a variety of indicators and for multiple years in the future (at least to 2050 and, for some models, until 2100). All new models will be open source and will be shared in GitHub. Their development will be documented in deliverable D6.7 – "Open-access models for case-study countries" and its update D6.8.

Funded by the European Union

2.2.2 Data inputs and outputs for models

Table 2 shows an indicative list of data inputs and outputs for models participating in this consortium, based on a similar list developed for the DMP of the PARIS REINFORCE project[5]. To assess national, regional, and global scenarios using energy system and sectoral models, it is necessary to go through an extensive data harmonisation process.

Table 2. Indicative inputs for harmonisation and output data

Data type	Category	Input/output
Population	Socioeconomic	Input
GDP	Socioeconomic	Input & output
Interest rates and exchange rates	Socioeconomic	Input
Labour participation and productivity	Socioeconomic	Input
Household size and space	Socioeconomic	Input
Physical activity data	Socioeconomic	Input
GHG emissions (CO_2, methane, others)	Emissions	Input & output
CO_2 emission factors	Emissions	Input
Installed capacity for different energy technologies and for storage	Energy	Input & output
Electricity and heat production mix	Energy	Input & output
Specifications of energy technologies	Energy	Input
Sectoral energy intensities	Energy	Output
Energy/electricity trade	Energy	Output
Food/agricultural demand	AFOLU	Output
Industrial technologies (Focus on steel sector) - specifications and availability	Industry	Input
Technological characteristics of vehicles	Transport	Input
Mobility characteristics and modal shifts	Transport	Output
Energy prices	Prices/costs	Input & output
Energy and power technology costs	Prices/costs	Input & output
Cost of technologies (all sectors)	Prices/costs	Input

Note: Data adapted from the DMP of PARIS REINFORCE.

[5] https://paris-reinforce.eu/sites/default/files/2022-12/D8.8%20Data%20Management%20Plan-Update%202_v1.00_SUBMITTED.pdf

Funded by the European Union

Harmonisation here refers to the process of aligning the inputs of different models to produce a model inter-comparison study. This helps to restrict model response heterogeneity mostly to the differences behind each model structure and theory. Socioeconomic parameters, techno-economic assumptions, policies and historical emissions are harmonised before the scenario runs. Data collected and processed for the harmonisation process adhere open access strategy of the project. Table 3 provides the list of the variables used for the harmonisation process.

Table 3. Harmonised variable data sources and format

Topic	Variable	Context	Recommended data	Data format
Socioeconomics	**Population**	EU27+Norway	Population data from Eurostat: Historical statistics through 2018 [2]; EUROPOP2019 projections from 2019 through 2100 [3,4]	csv
		OECD+7 except EU27/Norway	Population data from OECD Economic Outlook 109 Long-term baseline projections (available for 1990 through 2060), to ensure consistency with GDP projections [5]. Extend with growth rates from UN WPP2022 after 2060 if necessary (see below).	csv
		Rest of World	UN World Population Prospects 2022 [6] figures for all years (available from 1960 through 2100, with historical data until 2022)	csv
	GDP	All regions, through 2028	GDP from IMF World Economic Outlook April 2023 [7,8]8, in constant 2017 international dollars unless otherwise indicated by the research question	csv
		EU27+Norway, from 2029	Extend IMF forecast with linearly interpolated real GDP growth rates from the 2021 Ageing Report [9,10]	csv
		OECD+ except EU27/Norway, from 2029	Extend IMF forecast with real GDP growth rates from the OECD Economic Outlook 109 long-term baseline	csv
		Rest of World	Extend IMF forecast with GDP growth rates per working age capita from SSP2 (must be calculated using the projected population figures) multiplied by population growth rate from the harmonised population time series	csv
Techno-economics	Technology costs (for reference scenarios)	EU27 and comparable countries	Should be decided case-by-case, but most suitable option is likely to be technology costs from the EU Reference Scenario 2020[11,12].	csv
		Other regions, power-sector technologies	No firm recommendation at this point. Most suitable option is likely to be cost assumptions from IEA's World Energy Outlook 2022 [13]. Alternatives include technology costs from TIAM database as used in PARIS REINFORCE [14], or using the costs from the EU Reference Scenario 2020 in regions where this is appropriate, possibly adapting the costs levels if feasible.	csv
		Other regions, non-power technologies	No firm recommendations at this point. Options include the TIAM database (see above) or adapting costs from the EUR Reference Scenario 2020.	csv
	Fossil fuel prices	Historical prices	Regional prices from IEA datasets [15]. If participating modellers do not have access to proprietary IEA data, data for some years and regions can be extracted	csv

Funded by the European Union

			from the freely available World Energy Outlook 2022 report [13], or if global benchmarks are sufficient, some of these are available for free from the World Bank "pink sheet" [16]	
		Price projections	Regional price forecasts from the World Energy Outlook 2022 extended dataset[8], requires subscription. Participating modellers who do not have a license can extract some prices visually from charts in the World Energy Outlook 2022 report [13]. For long-term EU-specific exercises that do not need short-term trends, price projections from the EU Reference Scenario 2020 may be used [12].	csv
Energy	Energy production and consumption	Historical data11	IEA World Energy Balances (International Energy Agency, 2023e)	csv
Emissions	Energy-related CO_2, CH_4, N_2O	Historical data	IEA Greenhouse Gas Emissions from Energy dataset [17], if modellers have or can acquire access. Alternatively use EDGAR v7.0 [18] which is consistent with IEA but with less detailed breakdowns. Emissions can also be calculated from energy consumption data using default Tier 1 emission factors from the 2006 IPCC guidelines for GHG inventories, which are consistent with IEA emission factors.	csv
	IPPU CO_2 from cement	Historical data	Production data and emission factors from [19]. Updated data available on Zenodo [20].	csv
	F-gases and non-energy CO_2, CH_4, N_2O	Historical data	EDGAR v7.0 [18]	csv
	Other emissions	Historical data	CEDS [21]	csv
	Land-use change emissions	Historical data	No firm recommendation on harmonised data but should check that used or generated data falls within or close to range spanned by 3 bookkeeping models in the Global Carbon Budget 2022[22]. Use an average of the three models if a single harmonised dataset is required and no other constraints are implied by the research question. See subsection 3.7.5 for approach when alignment with national inventories is required.	csv
	All emissions	Infilling of historical and future emissions	Use Silicone software package to infill missing emission components, if required by the modelling exercise (e.g., for climate impact assessment) [23–25]	csv

For the stakeholder engagement, a database, managed by Brugel, of thousands of individuals across policy, academia, industry, and civil society will be created. The database is entirely GDPR-compliant and stored within Bruegel's Customer Relationship Management (CRM) software, Salesforce[6]. Contacts for IAM COMPACT will be informed of the project as soon as they are selected for involvement. Stakeholder details will be recorded in the CRM database, including helpful supplementary information to allow for effective categorisation of stakeholders.

[6] https://www.salesforce.com/eu/?ir=1

The full list of variables recorded for each stakeholder are listed in Table 4. The rich detail in the stakeholder pool will facilitate efficient targeting of stakeholders for the appropriate stakeholder engagement context.

Table 4. List of variables within stakeholder pool

Variable	Categories	Notes
Stakeholder Name	-	-
Organisation	-	-
Individuals' position within organisation	-	-
Stakeholder Category	• EU Policymaker • National government • International Institution • Private sector/industry • NGO • Labour/trade union • Academia • Other	-
Country	-	-
City	-	-
Partner Reference	-	Which partner made first contact with the stakeholder (if any)
Field of activity/sector	-	Specific field of activity within which the stakeholder operates
Level of activity	• Regional • National • Global	At which level does the organisation predominantly operate
Interest in engagement	• High • Average • Low	To the best understanding of the consortium partner (to be updated upon stakeholder feedback): how interested is this stakeholder in engaging with the project?
Influence (within level of activity)	• High • Average • Low	To the best understanding of the consortium partner (to be updated upon stakeholder feedback): what is the stakeholders' influence within their stated level of activity?
Email	-	-
Phone	-	-
Twitter	-	-
Focal point of contact	-	-
Date of engagement	-	-
Issues raised	-	Interests, motivations, strategies

Funded by the European Union

3 FAIR Data Guidelines

3.1 Making Data Findable

We will ensure that all project outputs are findable by using adequate identifiers and metadata. All deliverables and policy briefs will be uploaded in a Zenodo community that has been specifically created for the project[7] and where they will be automatically fitted with persistent digital object identifiers (DOIs). In the case of scientific publications, a DOI will be provided by the publishing journal, although we will still upload publications (or accepted manuscripts) in Zenodo and, potentially, also preprints. A versioning system will be also employed to track the state of each publication and ensure that interested parties can find the version they are looking for. For this, we will use Zenodo's paradigm where each new version gets a separate DOI, but a top-level DOI is also available, resolving to the latest version available. When each document is uploaded in Zenodo or on a journal, we will add adequate keywords to facilitate document retrieval via search engines. We will also use a consistent naming system for each type of publication:

- Scientific publications: "author(s)_name(s)_year" (e.g., Smith_et_al_2023.pdf)
- Policy briefs: "IAM_COMPACT_Policy_Brief_Title (e.g., IAM_COMPACT_Policy_Brief_IAMs_In_Policy.pdf)
- Deliverable: "IAM_COMPACT_DXX_Title" (e.g., IAM_COMPACT_D3.2_Open_Data_Management_Plan.pdf)
- Datasets: "lead_author_name_et_al_year_journal_dataset (e.g., Koutsandreas_et_al_2023_ESR_dataset)

Datasets of project outputs that underpin publications will also be archived in Zenodo and fitted with a DOI and adequate keywords. We will then link all datasets in Zenodo with the maDMP in ARGOS[8] and accompany them with rich metadata, using the Horizon Europe template provided in ARGOS (see Chapter 7 for more information). Similarly, all model code will be stored in public repositories in GitHub which will be then linked to Zenodo and the maDMP.

All model documentation, inputs, and outputs will be also published in the I^{2}AM PARIS modelling platform which will be continuously promoted as an integrated hub of information for the climate-economy and energy modelling community. For each dataset, we will include links to all related project publications in Zenodo and scientific journals to further increase findability. Finally, both the platform and the project website will include enough relevant content and an optimised sitemap structure to ensure findability in search engines like Google and Bing.

3.2 Making Data Openly Accessible

All project deliverables, policy briefs, training material and datasets will be published in Zenodo under Creative Commons licenses. In most cases, we will publish project outputs using the highly permissive CC BY license (version 4.0). Exceptionally, we may also consider less permissive licenses such as CC BY-NC to restrict commercial uses of project outcome, but only when this is necessary for specific and clear reasons by the author(s) of the publication or the dataset.

Similarly, scientific publications of the project will be published in journals offering open-access options in compliance with the Horizon Europe rules. When possible, we will publish in fully open-access journals, also considering the Open Research Europe publishing platform. In case that the available fully open-access options do not align with the scope of a publication, we will select a journal that offers a gold open-access option from the list of journals that the organisations of project partners have a publishing agreement with. In that way, we will ensure that all project insights become immediately available and exploitable by project audiences. In case that a project deliverable is linked to a scientific publication, we will still add the deliverable in Zenodo, but we may add an appropriate embargo period so that the deliverable becomes publicly available only after the paper is published.

[7] https://zenodo.org/communities/iam-compact/
[8] https://argos.openaire.eu/plans/overview/6510151d-9ab5-46d6-8468-388544c13b42

Funded by the European Union

All new models developed by the project will be published under open-source licenses. Table 5 shows all new models that will be developed in IAM COMPACT, along with the existing models that the new ones will be based on, their scope, and the partner that will develop them. It is noted that these licenses refer to the model code and not the input data. The selection of licences for both model code and input data will be finalised during the project.

Table 5. Preliminary selection of open-source licenses to use for new model code

Country	Resource	Scope	Partners	License	Links
Ethiopia	OnSSET	Electrification model	AAiT, KTH	MIT	https://github.com/OnSSET/onsset
	MicrogridsPy	Microgrids sizing	Polimi	EUPL-1.1	https://github.com/SESAM-Polimi/MicroGridsPy-SESAM
	EnergyPlan	Energy system (operation)	Aalborg		https://www.i2am-paris.eu/detailed_model_doc/energyplan
Kenya	CLEWs	Climate, Land, Energy, Water systems model	TUM, KTH	MIT	https://github.com/OSeMOSYS/CLEWs
Ukraine	OSeMOSYS	Energy system	KEI	Apache 2.0	https://github.com/OSeMOSYS/OSeMOSYS
	EnergyPlan	Energy system (operation)	Aalborg		https://www.i2am-paris.eu/detailed_model_doc/energyplan
Sri Lanka	CLEWs	Climate, Land, Energy, Water systems model	RJRT, KTH	MIT	https://github.com/OSeMOSYS/CLEWs
	MARIO	Input-Output macroeconomic modelling framework	Polimi	GPL-3.0	https://github.com/it-is-me-mario/MARIO

In terms of file formats, we will strive to use well-known formats that can be opened by freeware software such as pdf, csv, txt, mp3, and mp4 files. We will strive to avoid sharing content through proprietary file formats such as docx and xlsx. When this is not feasible (for instance, when we need to publish spreadsheet files with multiple tabs), we will accompany files in proprietary formats with links to compatible freeware software such as the Open Office suite.

We will also ensure that all open project outputs will remain available for as long as possible. All project publications and datasets will be published in established online repositories such as Zenodo and GitHub where high availability is expected for many years to come. This is especially the case for Zenodo where its operation is supported by the EU. The project website will be also kept online for at least three years after the project's end to support the dissemination and findability of project outcomes. Project coordinator NTUA will also ensure the longevity of the I^{2}AM PARIS platform for at least four years after IAM COMPACT ends (till around 2030). The partners will also undertake efforts to ensure the platform's sustainability, either by future projects that will take over its operation or through appropriate funding mechanisms. The Horizon Results Booster service of the EU will be also consulted to find these mechanisms.

3.3 Making Data Interoperable

We will achieve high interoperability of project data by using adequate data formats and by providing informative metadata. As suggested in Section 3.2, all project datasets and reports of the project will be shared through widespread formats such as pdf, csv, and txt, avoiding proprietary formats when possible. The documentation of

Funded by the European Union

all new models along with the results of the scenario analysis will be formatted based on the IPCC AR6 reporting templates, ensuring that the wider climate-economy modelling community can use them, while also achieving interoperability with other relevant platforms of the community such as the Scenario Explorers by IIASA. Metadata for all project datasets will be added in IAM COMPACT's maDMP in ARGOS, using the format template of Horizon Europe. As the maDMP is machine actionable, it will be fairly easy to read the internal representation of metadata and potentially convert it to another format, further ensuring the interoperability of project datasets.

3.4 Making Data Reusable

As suggested in the previous section on open access, by releasing all deliverables and datasets through CC BY license we will also support their uptake by interested parties. Similarly, all scientific papers will be published under open-access licenses and will be made available directly after acceptance by the journals. We will use more restrictive licenses, e.g., including clauses to limit commercial use, only in exceptional cases and when it is fully necessary. We will also ensure the reusability of datasets by using the Horizon Europe metadata scheme in the maDMP of the project. For each dataset, the scheme will provide a short description of the data, links with publications and other datasets, and guidelines for the specific dataset related to FAIR practices, allocation of resources, and security and ethical aspects. Lastly, all modelling documentation and results will be formatted using the reporting templates of IPCC AR6 to ensure reusability by the wider modelling community of integrated assessment and climate change mitigation.

Funded by
the European Union

4 Data Security

NTUA and Imperial will overview all data management in the project and will ensure the security of all project data through robust data storage and secure platforms for communication and data exchange. On the latter, NTUA has created a dedicated workspace in its enterprise version of Microsoft Teams as an internal communication system for video-calls and chatting among project partners (see also Milestone 2). This system is also connected to a secure instance of Microsoft SharePoint which will serve as the exclusive data exchanging system for the project. The management and security of these systems is supported by the admins and the Data Protection Officer (DPO) of NTUA, while their servers are within the EU (Greece), ensuring compliance with GDPR and relevant EU legislation. This will be important as SharePoint will be also used to store contact details of project stakeholders that will need to be fully secure. Similarly, the personal data of newsletter subscribers will be stored in the MailerLite account of NTUA which is also GDPR-compliant.

Apart from the SharePoint, other data storage systems used in the project include the databases of the project website and the I^2AM PARIS platform. For both databases, NTUA have implemented disaster recovery and backup policies to ensure that the data is safe from loss caused by a disaster such as a critical systems failure, fire, theft, or natural disaster. A similar process is followed by NTUA for the SharePoint system while there is also a versioning system in place that protects the users from accidentally deleting or modifying data.

The project's communities in Zenodo and GitHub will be also used to store data during the process, and, most importantly, to preserve all created datasets and publications after the end of the project. The possibility of data loss in these repositories is very low, as all files and documents are stored in multiple online servers that ensure redundancy. Additionally, it is highly unlikely that these repositories will close operations and, even in that case, they will migrate all content to suitable archives such as the servers of the Software Heritage Foundation and Internet Archive.

Funded by
the European Union

5 Allocation of Resources

Most of the practices described in Section 3 for ensuring FAIR data in the project are not requiring any costs from the project. All deliverables and datasets will be uploaded in Zenodo which is free to use, and we will also use the free version of GitHub to store project code. Similarly, the documentation of datasets in the maDMP in ARGOS is also free of charge.

Other activities for ensuring FAIR data require resources which have been considered in the project's Grant Agreement. The extension and hosting of I²AM PARIS requires funds that have been budgeted under WP3, while funds on the development and maintenance of the project website have been considered in the budget of WP1 and the defined purchase costs for NTUA. We have also earmarked a part of the budget for publishing in open access journals. Most of this budget is managed by the project coordinator NTUA while all partners have some funds available for individual open access publications related to their work in the project. Suggested options for publishing in open access journals are shown in Table 6.

Table 6. Suggested options for open-access scientific publishing in IAM COMPACT

Access type	Funder	Fees	License
Publish in a fully Open Access journal	Paid through IAM COMPACT's Grant	Article Processing Charges indicatively ranging between 130€ and 9,500€	CC BY 4.0 CC BY-NC-ND 4.0
Publish in a journal that has the option of Gold Open Access	Paid through publishing agreements between the organisations of project partners and the publisher		

In terms of responsibilities of project partners, NTUA and Imperial will have the overview of all data management in the project. The same partners will ensure that all project data and publications are uploaded in Zenodo and linked to the project website, the maDMP, and the I²AM PARIS platform. The platform will be further developed by NTUA, including creating new visualisations and applications based on the project's data. All project partners will be responsible for correct data handling and curation based on the guidelines of the DMP, including that model code is frequently uploaded in the IAM COMPACT community in GitHub. As also mentioned above, NTUA will be responsible for keeping the website and the platform online and all related project data available for at least three years after the end of the project. As far as the platform I concerned, NTUA are exploring and securing ways to further extend its lifetime, notably through its continued use in ongoing (DIAMOND[9], ENCLUDE[10]) and new (TRANSIENCE, EU-CHINA-BRIDGE) Horizon Europe projects.

[9] https://climate-diamond.eu/
[10] https://encludeproject.eu/

Funded by the European Union

6 Ethical Aspects

All data collection and management activities of the project will be compliant with the EU GDPR regulation and with the national privacy and data protection laws of the countries of each partner. For most activities related to model development, there are not ethical or legal aspects apart from respecting the licenses of databases that will be used as sources of modelling inputs. In contrast, ethical aspects become especially important in all the co-creation activities of the project (WP2), where we will collect the perspectives of different project stakeholders through workshops, interviews, and surveys. For the workshops we will use the Chatham House Rule ("when a meeting, or part thereof, is held under the Chatham House Rule[11], participants are free to use the information received, but neither the identity nor the affiliation of the speaker(s), nor that of any other participant, may be revealed."). Minutes of the workshops will heed this rule and will avoid linking individuals to specific statements. The same process will be used for documenting interviews, while surveys will also avoid questions on personal details other than those that are needed for research reasons (e.g., whether a respondent is from academia or policymaking). In all engagement activities, we will ask for an explicit and clear informed consent from the participants. The form used to get this informed consent and further details on ethical aspects of the stakeholder engagement activities of the project will be included within the deliverables of WP2.

As mentioned in Section 5, all contact details and feedback of project stakeholders will be stored in the project's secure SharePoint instance while the contact details of the newsletter subscribers will be stored in the MailerLite account of NTUA. Both the SharePoint and the MailerLite instances are hosted within the EU and are GDPR compliant. We will avoid exchanging contact details en masse through unsecure channels such as emails. GDPR rules would be used in all personal and sensitive data transfers between project partners.

[11] https://www.chathamhouse.org/about-us/chatham-house-rule

Funded by
the European Union

7 Machine actionable DMP in Argos

In parallel with this report, the IAM COMPACT maDMP in ARGOS will provide a detailed overview of all datasets generated, curated, or managed during the project[12]. The maDMP will be updated whenever a project dataset is created or modified during the project and a summary of its contents will be provided in the update of this report (D3.3). Personnel from NTUA and Imperial will be responsible for creating entries for new datasets in the maDMP while all project partners will be responsible for checking that the metadata provided for their datasets in the maDMP is correct. The following process will be used for creating a new dataset in maDMP:

1. The project partner(s) that created the dataset (henceforth called "data creators") will share it with NTUA and Imperial ("data managers").
2. The data managers will then upload the dataset on Zenodo.
3. The data managers will also create a new dataset in the maDMP and prefill it by searching the name of the dataset through the search engine of ARGOS.
4. The data managers will complete the metadata for the dataset based on guidance from this report and its updates and will also link the dataset to relevant deliverables or scientific publications.
5. The data creators will be invited to check these metadata and ensure their accuracy.
6. The data managers will update the maDMP with the new dataset.

A similar process will be used for new model code, where links will be created between the model repositories in GitHub and Zenodo to ensure that the code base of the models receives a unique DOI. This Zenodo entry will be then linked to the maDMP which will be documented with adequate metadata as above. All these processes will be evaluated based on the experience of data managers and creators and may be adapted and optimised further during the duration of the project.

[12] https://argos.openaire.eu/plans/overview/6510151d-9ab5-46d6-8468-388544c13b42

Funded by
the European Union

Bibliography

[1] Lin D, Crabtree J, Dillo I, Downs RR, Edmunds R, Giaretta D, et al. The TRUST Principles for digital repositories. Scientific Data 2020 7:1 2020;7:1–5. https://doi.org/10.1038/s41597-020-0486-7.
[2] Eurostat. Demography, population stock and balance (demo) [Data set] n.d. https://ec.europa.eu/eurostat/databrowser/explore/all/popul?lang=en&subtheme=demo (accessed November 30, 2023).
[3] Eurostat. EUROPOP2019 - Population projections at national level (2019-2100) (proj_19n; Short-term update 2022-09-28) [Data set] 2022. https://ec.europa.eu/eurostat/databrowser/explore/all/popul?lang=en&subtheme=proj.proj_19n (accessed November 30, 2023).
[4] Eurostat. EUROPOP2019—Population projections at regional level (2019-2100) (Short-term update 2022-09-28) [Data set]. 2022. https://ec.europa.eu/eurostat/databrowser/explore/all/popul?lang=en&subtheme=proj.proj_19r (accessed November 30, 2023).
[5] OECD. Economic Outlook No 109—October 2021 - Long-term baseline projections (EO109_LTB) [Data set]. 2021. https://stats.oecd.org/Index.aspx?DataSetCode=EO109_LTB (accessed November 30, 2023).
[6] United Nations. World Population Prospects 2022 [Data set] 2022. https://population.un.org/wpp/ (accessed November 30, 2023).
[7] International Energy Agency. World Energy Outlook 2022 Free Dataset [Data set]. 2023 n.d. https://www.iea.org/data-and-statistics/data-product/world-energy-outlook-2022-free-dataset (accessed November 30, 2023).
[8] International Energy Agency. World Energy Outlook 2022 Extended Dataset [Data set]. 2023 n.d. https://www.iea.org/data-and-statistics/data-product/world-energy-outlook-2022-extended-dataset (accessed November 30, 2023).
[9] European Commission D-G for E and FA. The 2021 ageing report: Economic & budgetary projections for the EU Member States (2019 2070) 2021. https://op.europa.eu/en/publication-detail/-/publication/8b1015a6-ead6-11eb-93a8-01aa75ed71a1/language-en (accessed November 30, 2023).
[10] European Commission D-G for E and FA. Ageing Report 2021 Data [Data set] 2021. https://data.europa.eu/data/datasets/ageing-report-2018?locale=en (accessed November 30, 2023).
[11] European Commission. EU Reference Scenario 2020. N.d n.d. https://energy.ec.europa.eu/data-and-analysis/energy-modelling/eu-reference-scenario-2020_en (accessed November 30, 2023).
[12] European Commission D-G for CA, De Vita A, Capros P, Paroussos L, et al. EU reference scenario 2020 : energy, transport and GHG emissions : trends to 2050. 2021. https://doi.org/10.2833/35750.
[13] International Energy Agency. World Energy Outlook 2022 2022. https://www.iea.org/reports/world-energy-outlook-2022 (accessed November 30, 2023).
[14] Napp TA, Few S, Sood A, Bernie D, Hawkes A, Gambhir A. The role of advanced demand-sector technologies and energy demand reduction in achieving ambitious carbon budgets. Appl Energy 2019;238:351–67. https://doi.org/10.1016/J.APENERGY.2019.01.033.
[15] International Energy Agency. Energy Prices [Data set] 2023. https://www.iea.org/data-and-statistics/data-product/energy-prices (accessed November 30, 2023).
[16] World Bank. Commodity Markets 2023. https://www.worldbank.org/en/research/commodity-markets (accessed November 30, 2023).
[17] International Energy Agency. Greenhouse Gas Emissions from Energy [Data set] 2023. https://www.iea.org/data-and-statistics/data-product/greenhouse-gas-emissions-from-energy (accessed November 30, 2023).
[18] Branco A, Crippa M, Guizzardi D, Banja M, Solazzo E, Muntean M, et al. Emissions Database for Global Atmospheric Research (v7.0_FT_2021) [Data set]. European Commission, Joint Research Centre (JRC) 2022.
[19] Andrew RM. Global CO2 emissions from cement production. Earth Syst Sci Data 2018;10:195–217. https://doi.org/10.5194/ESSD-10-195-2018.
[20] Andrew R. Global CO2 emissions from cement production (Version 230428) [Data set] 2023. https://doi.org/10.5281/ZENODO.7875557.
[21] O'Rourke PR, Smith SJ, Mott A, Ahsan H, McDuffie EE, Crippa M, et al. CEDS GitHub repository. Joint Global Change Research Institute 2021. https://github.com/JGCRI/CEDS (accessed November 30, 2023).
[22] Friedlingstein P, O'sullivan M, Jones MW, Andrew RM, Gregor L, Hauck J, et al. Global Carbon Project. (2022). Supplemental data of Global Carbon Budget 2022 (Version 1.0) [Data set]. Earth Syst Sci Data 2022;14:4811–900. https://doi.org/10.5194/ESSD-14-4811-2022.

Funded by
the European Union

[23] Lamboll RD, Nicholls ZRJ, Kikstra JS, Meinshausen M, Rogelj J. Silicone v1.0.0: An open-source Python package for inferring missing emissions data for climate change research. Geosci Model Dev 2020;13:5259–75. https://doi.org/10.5194/GMD-13-5259-2020.
[24] Lamboll RD, Nicholls Z, Kikstra J. Silicone GitHub repository 2022. https://github.com/GranthamImperial/silicone (accessed November 30, 2023).
[25] Lamboll RD, Nicholls Z, Kikstra J. Silicone documentation. ReadthedocsIo 2022. https://silicone.readthedocs.io/en/latest/search.html (accessed November 30, 2023).

Funded by
the European Union