**Supplementary Online Material (SOM)**

[**SOM1. Energy modelling for policy support artifacts**](#_pbrdf1dy2t6j) **16**

[**SOM2. Selected communities and energy policy**](#_dx2rgzsmh27u) **17**

[**SOM3. Selected Incentives and U4RIA needs by community or role**](#_mtiooju1ktd3) **18**

[**SOM4. Strengths and weaknesses of insourcing and outsourcing**](#_j553p6fz637x) **19**

[**SOM5. Auditability**](#_bn2xz0guiczg) **20**

[**SOM6. Open licensing of data and code**](#_p6elwr6075jt) **21**

[**SOM7. Template of Annex to EMoPS Terms of Reference**](#_14xq1jn0lavo) **24**

# SOM1. Energy modelling for policy support artifacts

This section outlines elements of a stylized national energy modelling process that should ideally carry on despite political changes. Derived from Mirakyan and De Guio (2014),[2](https://www.zotero.org/google-docs/?bv8Rtx) it typically involves an orientation,[1](https://www.zotero.org/google-docs/?MDOdhO) an analysis and an insight generation phase involving several actors and techniques.[43](https://www.zotero.org/google-docs/?W0kbBX) It does so by developing scenarios that are part of an analytical construct to deliver insights. (Even if the output is a quantitative ‘cost-benefit-analysis’, this typically involves the development of a ‘reference’ or ‘counterfactual’ scenario without the policy intervention, and is compared to a scenario with the policy intervention.[44](https://www.zotero.org/google-docs/?uyZMdM) And, of course there are more complex efforts that involve assessment of robustness[45](https://www.zotero.org/google-docs/?jPdNEH), deep uncertainty,[46](https://www.zotero.org/google-docs/?9B6Iej) futures,[47](https://www.zotero.org/google-docs/?rrUGIn) etc, each with different strengths and weaknesses.[48](https://www.zotero.org/google-docs/?rUTrIL))

To get from data to scenarios to insights, there are digital and organisational workflows. An organisational workflow describes the organization of actors. Those actors in communities each having varying levels of participation in the process. There is often no tightly predefined digital or community/organisational workflow. They are often defined in situ, being both context and constraint specific.[1](https://www.zotero.org/google-docs/?PDkZDU) The figure is of course a simplification (with variations based on context[49](https://www.zotero.org/google-docs/?uhu1Vs)). The process is often iterative, and fluid. (This is not to be mistaken by a linear ‘rational’ approach the figure might imply!)

**Diagram

Description automatically generated with low confidence**

**Figure SOM1. A simplified stylised linear physical and thought process of energy modelling for policy support. In reality that process can be far more interlinked, evolve and iterate. Adapted from**[**2**](https://www.zotero.org/google-docs/?TlY0dn)**. (Note that communities in grey can take on a greater or lesser role and be differently composed depending on the context).**

# SOM2. Selected communities and energy policy

In figure SOM2 we extend work by Sharpe and Victor (2019)[50](https://www.zotero.org/google-docs/?ULrEgA) and summarise key communities affected by configurations of the energy system. The same might be engaged in the EMoPS proces, in particular via its scenarios.

Diagram

Description automatically generated

**Figure SOM2. Communities affected by the configuration of the energy system from policies that may be influenced by EMoPS.**

# SOM3. Selected Incentives and U4RIA needs by community or role

|  |  |  |
| --- | --- | --- |
| **Community** | **Incentive** | **Needs** |
| Modelling community | Substantial gains from a collaborative approach to domain‑wide issues, many such issues revolve around data collection, curation, and transformation. |  |
| Modelling team | Performing rapidly and skillfully to meet demand, the most obvious of which are the policy insights needed. (Due to the degrees of freedom, imperfect information, and limited time, the analysis can always be tweaked. This has led some to call EMoPS an art not a science[51](https://www.zotero.org/google-docs/?w1Nt04).)  These ‘tweaks’ can be the focus of the analyst time, rather than the perceived mundainess of documenting, curating and archiving: the digital and organisational process, data sources, scenario design, references references, workflow descriptions and ensuring easy retrieval etc. | Clearly articulated terms of reference that ensure appropriate documentation, filing, testing and retrieval of all EMoPS elements, immediately this is for knowledge management and the operation of the modelling team. Ultimately it is to allow for eMOP audit.  The extra terms, however, should be carefully crafted to ensure minimum interference with the delivery of the policy support required. While maximising the useful impact they bring in terms of better analytics and good governance practice. |
| Decision or policy maker | To receive a clear input, a clear answer to his/her question, and in a short amount of time. A good EMoPS product provides answers to the decision-makers questions or a set of clear options to make a decision.  In so doing, decision-makers need justifiable information and process to legitimise the insights, and policy that is supported. After the fact, some of those insights will be shown to be wrong due to the plethora of deep unknowns at the time of the EMoPS. However, if the eMOPS process demonstrably meets appropriate due diligence standards, the decision maker could perhaps be shown to have acted appropriately.. | An auditable set ofEMoPS elements need to be delivered along with the policy insights that it produces.  Of special interest will be the overall analytical and organisational design was appropriate.  The design and EMoPS elements should meet good analytical and governance practice. |
| Coordination group | The function of the coordination group differs by its context and constitution however, there are typically important common goals therein, including:   * Ensuring coherence of information flow and data generation. * Building consensus * Coraling appropriate participation from key stakeholders to the public at large | A fit for purpose organisational workflow needs to be developed that allows for a mapping of stakeholders and their roles.  This will require EMoPS elements developed in such a way that they facilitate appropriate suitably documented participation. |
| National or International planning agencies. | To improve national (or partner) development by improving the evidence base for policy making. | Of importance will be to show value for money and accountability in spending. Among other things, that implies specifying to an appropriate, ‘tangible’ and auditable product. |
| Stakeholders, civil society and other constituencies | It is vital that the broader society is offered engagement and consultation and that such processes are transparent and accountable. | This will require an auditable process and appropriate organizational workflow. With clear opportunities and entry points for meaningful participation.  Depending on the context this might vary from simply knowing the process is auditable, to the co-creation of scenario building and building national capacity. |

Table SOM3. Selected communities incentives and needs.

# SOM4. Strengths and weaknesses of insourcing and outsourcing

Focusing on the modelling team, patterns of its organisations are recently documented.[52](https://www.zotero.org/google-docs/?lAin2G) The International Renewable Energy Agency’s (IRENA) campaign on Long-term Energy Scenarios for the Clean Energy Transition [22](https://www.zotero.org/google-docs/?oVHtN5) has identified several key factors that broadly determine the approach to source the EMoPS modelling team in government while producing long-term energy transition scenarios (LTES) from modelling efforts; these are summarised in figure 2.

The relative merits of in-sourcing and out-sourcing are not always clear‑cut. The German government, for instance, has no in‑house LTES capacity but instead funds a network of research institutes to undertake the analysis it needs..

Diagram

Description automatically generated

**Figure SOM4. Advantages and challenges of insourcing and outsourcing EMoPS scenarios. Source IRENA**[**52**](https://www.zotero.org/google-docs/?RyD8pm)

# SOM5. Auditability

The results of poor U4RIA will affect the ability to test and audit outputs. The potential consequences are serious, and noted earlier. Further, transparency, openness, accountability and thus auditability may also be needed to release funding, underpin transactions, enter into negotiations or motivate future energy sector actions. In fact, auditability will likely be a key theme for all countries reporting under the UNFCCC’s Enhanced Transparency Framework.[53](https://www.zotero.org/google-docs/?MSohln) In order to audit EMoPS we suggest that to the auditors, all elements of EMoPS need to be U4RIA compliant. A note on the meta-data and licensing requirements needed to enable that are included in figure SOM5.

Diagram

Description automatically generated with medium confidence

**Figure SOM5. Metadata and licencing requirements allowing for auditability**

# SOM6. Open licensing of data and code

The use of open licenses enables many of the processes described in this document. Without suitable open licenses that permit both use and reuse, users cannot lawfully run third‑party computer models, nor can they be certain that they can legally modify and share third‑party datasets[54](https://www.zotero.org/google-docs/?nO8v9X). These and related use‑cases fall under intellectual property (IP) law. Open licenses give recipients explicit permissions to study, use, improve, and share the work in question without first having to defer to potential rights holders. The most well‑known open licenses are those from Creative Commons (CC) — while noting that not all CC licenses class as open as just defined. Three types of work are relevant to public policy analysis: computer programs, datasets, and documentation. Each is handled quite differently under IP law and therefore also covered separately here. Computer programs are automatically protected by copyright and users need a suitable license to simply run a program, let alone modify and share it with others[55](https://www.zotero.org/google-docs/?uext5v). The legal definition of a computer program is broad and includes small scripts written to process or visualize datasets. Computer programs require dedicated licenses because there are a number of technical and legal issues to consider. These licenses also vary as to whether they allow reuse in proprietary software (so‑called permissive licenses) or not (so‑called reciprocal or copyleft licenses). The choice in this regard is for the developers concerned and does not impact materially on issues discussed in this report[56](https://www.zotero.org/google-docs/?jCvLfI).

Data generally and datasets specifically are less well resolved under IP law and it is often not clear whether IP rights even apply. For example, machine‑generated datasets may well not attract copyright but that view has not been tested in law[57](https://www.zotero.org/google-docs/?jwqe5O). Like software, datasets also require dedicated licenses because there are specialized legal issues to traverse[58](https://www.zotero.org/google-docs/?I0j2Rz). Irrespective, the solution is for data providers to add either a Creative Commons CC‑BY‑4.0 attribution license or a Creative Commons CC0‑1.0 public domain waiver, prior to publication and at their discretion[59,60](https://www.zotero.org/google-docs/?iNV0Hj). Documentation, for our purposes, is presumed to include tables containing numeric data and diagrams depicting that data or other information. Tables are therefore a form of dataset with the same IP issues, which means that data‑capable licensing should also be used for documentation. (it should be noted that tables and diagrams can be licensed separately but that approach is not often used). Metadata should be licensed CC0‑1.0 in all cases.

Table SOM6: Recommended open licenses by type of work. Their SPDX license abbreviations are used where appropriate.

|  |  |  |
| --- | --- | --- |
| Type of work | Recommended licenses | Comment |
| computer program | MIT, Apache‑2.0, GPL family | required to lawfully run a computer program |
| dataset or database | CC‑BY‑4.0, CC0‑1.0 | often unclear whether intellectual property applies, open licensing removes that uncertainty |
| documentation | CC‑BY‑4.0, CC0‑1.0 | recommendations presume tables and diagrams are present |

Table SOM6 summarizes the preceding discussion. Open licenses remain central to reuse. Open licensing removes the need to determine which IP rights apply or not and thereby provides users with legal certainty. Without suitable open licenses, much of what is discussed in this report simply cannot occur.

A screenshot of a computer

Description automatically generated with low confidence

**Figure SOM6. Open licence selection suggestions**

# SOM7. Template of Annex to EMoPS Terms of Reference

**Special requirements for data, metadata and models**

These are general requirements for open access data and energy modelling assignments as prepared by the [61](https://www.zotero.org/google-docs/?UMvPKm) in compliance to its key Principles and strongly recommended to be included in ToRs where there will be significant data collection and modelling.

**Property and Confidentiality of Data and Modelling Outputs**

Data collected and deliverables produced under this TOR – including metadata, intermediate data and data collection and analysis methodologies – are the property of [“PROJECT OWNER”][[1]](#footnote-1) and considered confidential information. The Consultant or vendor will protect the confidentiality of establishments and individuals participating in the provision of data or information at all stages. Exception to such protection of confidentiality is at the sole determination of [“PROJECT OWNER”], provided such an exception is allowed under applicable national laws. The Consultant or vendor will ensure that no data or related documentation collected or compiled under these TORs are distributed for commercial or non-commercial purposes to third parties, nor will they be used by the Consultant, firm, its staff or sub-contractors for purposes other than those expressly stated in these TORs, without the prior written approval of the [“PROJECT OWNER”].

**Compliance with the U4RIA data and modelling principles**

[“PROJECT OWNER”] has been part of the process to define a set of principles to promote the robust, accessible and transparent delivery of energy modelling for policy support. Therefore, **all data collected and deliverables produced under this TOR must** **comply with the principles of Ubuntu, Retrievability, Repeatability, Reconstructability, Interoperability and Auditability (U4RIA)**, as illustrated below.

***Ubuntu (Community Engagement)***

Under the provision of this TOR, the Consultant agrees, to the extent specifically agreed with [“PROJECT OWNER”], **to** **engage with the national and international energy planning and modelling community** **about the outputs and methodology produced under this TOR in one or more of the following ways:**

* **Peer review:** The output data and methodology used are reviewed by a group of national and international experts. Indication of the peer reviewers’ names and the peer review process (including the comments received) must be documented and submitted to [“PROJECT OWNER”].
* **Presentation of final deliverables:** The output data and methodology used are presented to a list of relevant national and international stakeholders agreed with [“PROJECT OWNER”].
* **Internal capacity building:** The Consultant builds the capacity of the main governmental end users in order to transfer the knowledge and ownership of the output data and methodology used. This should also involve regular interaction with the main end users throughout the modelling process.
* **External capacity building:** The Consultant builds the capacity of a list of relevant national and international stakeholders agreed with [“PROJECT OWNER”], in order to transfer the knowledge and ownership of the output data and methodology used.

***Retrievability***

**In accordance with the “**[**Key principles for improving the support to strategic energy planning in developing and emerging economies**](https://energyeconomicgrowth.org/index.php/publication/key-principles-improving-support-strategic-energy-planning-developing-and-emerging)**”, [“PROJECT OWNER”]** **intends to make all data and other deliverables produced under these TOR publicly available,** unless [“PROJECT OWNER”] believes that the public dissemination of the data will violate confidential information.

In particular, [“PROJECT OWNER”] intends to upload all data and other deliverables produced under these TOR on the following open access platforms and websites:

* [LIST OF WEBSITES AND PLATFORMS]

The Consultant will provide advice to [“PROJECT OWNER”] on options for uploading all data and other deliverables produced under this TOR.

***Repeatability***

Essential metadata describing, *inter alia*, data in and out, model generators, model generated, processes followed, ‘storage’ of that information and related meta-information should be provided for all data products. Core metadata requirements for each data product are:

* Names and contacts of the authors/consultants and lead institution
* Problem or policy issue analysed
* Type of modelling framework (e.g. accounting, optimisation, simulation, etc.)
* Version of the software considered
* Available code base / instance of both the:
  + Software i.e. model generators (e.g. MAED, LEAP, OSeMOSYS, MESSAGE, MARKAL, MAPS, etc.) and
  + The country specific model ‘generated’ and ‘calibrated’
* Input data (plus related meta-information) to the model, scenario assumptions and the outputs obtained
* Techno-economic information and degree of detail of the energy system structure represented in the model (what components/technologies included and how they are interlinked)
* System boundaries and restrictions (technical, environmental, social) applied and why
* If applicable, policies evaluated, results interpreted, and policies formulated based on the results
* Sensitivity and uncertainty analyses carried out.

***Reusability***

It is important that the modelling outputs from this assignment build on previous relevant modelling efforts and can contribute to future modelling efforts. In this regard,

* on the one hand, [“PROJECT OWNER”] commits to share with the Consultant any relevant material in its knowledge and facilitate the Consultant’s retrieval of previous relevant modelling outputs;
* on the other hand, the Consultant must document its efforts to retrieve and build on relevant modelling outputs, including by submitting a list of stakeholders contacted and data / material received.

In addition, the data and deliverables produced under this TOR should be provided in formats that allow as much as possible to be the basis for future modelling efforts. Therefore, data should be provided to [“PROJECT OWNER”] in at least one machine readable, non-proprietary open file format that complies with the Open Definition. Conformity with current standards and practices, including ontologies and metadata as well as tabular formats such as CSV and tab-delimited text, or geospatial formats such as Shapefile or GeoJSON satisfy this requirement. (PowerSystems.jl which embeds a complete and coherent representation of the system and can service model-specific needs represent an example of the former.) Excel, STATA, or other proprietary data formats may optionally be used in addition to at least one open format. PDF and Word are not acceptable formats for data.

Data should be provided according to recognized standards and encodings whenever possible. Data standards are available for many types of data; for instance, GTFS for transport data, or Data Document Initiative for metadata. The Open Geospatial Consortium documents standards for a broad range of applications and disciplines.

***Reconstructability***

The Consultant should provide a clear description of the workflows to move from the modelling input to the output data, so that the process to obtain the output can be reconstructed by a third party. As a minimum, this will include:

* Reference(s) of the original input data, e.g. source name and author, publication date etc. – or no reference, i.e. the value is an assumption from the modeller
* Level of manipulation (e.g. single value from one source, calculation from multiple values from one source, calculation from multiple values from multiple sources)
* Type of manipulation (e.g. average, mean, straight interpolation etc.)
* Time series included, i.e. details of the years the final value refers to
* Further comments (including description of assumptions involved).

***Interoperability***

The modelling output data from this assignment should be delivered in a form that is conducive for their utilisation by other models with minimal manipulation. To achieve acceptable levels of interoperability, as a minimum, the Consultant commits to:

* Submit well documented or annotated copies of base data to be appropriately stored by [“PROJECT OWNER”]
* Ensure it is on accessible media
* Ensure that the required model version can be downloaded
* Made the data compliant with the Standard Interchange Formats [NAME AND LINK / REFERENCE OF THE SPECIFIC STANDARD INTERCHANGE FORMAT].

***Auditability***

It is important that all the previous U4RI principles are followed, so that a successful audit of the data and deliverables produced under this TOR can be carried out. [“PROJECT OWNER”] reserves the right to include the compliance to the U4RIA principles as described in this Annex in an official audit and link the full or partial release of funding under this contract to the audit’s outcomes.[l12]

1. Please change everywhere in the document the reference to “[“PROJECT OWNER”]” with the appropriate wording of the project outputs’ “owner”, according to your organisation and project’s needs. Depending on the specificities of the case, this may be the Donor / Development Partner (e.g. UK FCDO, IRENA), the main government partner (Ministry of Energy), a third party or even a combination of those options. [↑](#footnote-ref-1)