

Remarks on the PRIMES databank

To: The European Commission
Regarding: EU Reference Scenario 2025 — Stakeholder Consultation on Technology Assumptions
Web reference: <https://ec.europa.eu/eusurvey/runner/scenario2025>
Email: EC-TECHNOLOGY-ASSUMPTIONS-2025@ec.europa.eu
Submission deadline: 30 September 2024

From: Robbie Morrison
Address: Schillerstraße 85, 10627 Berlin, Germany
Email: robbie.morrison@posteo.de

Date: 29 September 2024
Release number: **01**

Preamble

The European Commission is currently consulting on the technological assumptions used in the PRIMES energy system model. This submission offers some comments.

Robbie Morrison has been involved in high-resolution national energy system modeling since 1995, open source development of same since 2003, and the provisions of genuinely open data in support of transparent modeling and open science since 2018. Robbie contributes to the Open Energy Modelling Initiative community and is the lead administrator for their online discussion forum.

This submission is provided in a personal capacity.

Introduction

My comments relate to the structuring and the technical and legal reusability of the information provided in the PRIMES spreadsheets.

My comments do not relate to the parameters selected, the values assigned, nor the sources cited underpinning the given information in general terms. Nor am I considering whether the information used carries over from earlier databases or has been updated for this release. In short, my comments do not relate to the specifics of the data provided.

To highlight the legal reusability issue: I ask could I modify a data pair from one of the spreadsheets and republish that modified spreadsheet for the benefit of the open energy modeling community. The answer is no — or at least not without some considerable legal risk.

Indeed, the spreadsheets are notated: "© E3-Modelling - Confidential - Not to be used or further distributed without permission". I also note that "use" is not a term-of-art under copyright law and its legal interpretation remains unclear at best.

In order to ensure technical comprehensibility, the underlying data model should be articulated. Knowledge of the data model would also assist transparently and reuse

Information provenance should act at a granular level too. The would allow independent researchers to understand the values provided, their source, and their likely integrity,

Something more than a set of spreadsheets and an accompanying commentary is needed.

I will argue in this submission:

- that the underpinning data model be published
- that more sophisticated data management is required and will advocate the use of data mesh concepts
- that the legal provenance and public licensing terms, if any, be recorded

In short, the data provided needs to include information that will make that data fully understandable and reusable by third party modelers. In this context, simply publishing the spreadsheets might meet a very low-bar definition of transparency but does nothing for open science and independent analysis. Indeed the current practice falls well short when judged by current good practice standards for research data management.

My submission advocates a more sophisticated approach and is divided into three interconnected parts:

- the need to articulate the underlying data model
- provision of metadata detailing upstream provenance, technical attributes, and internal changes
- that same metadata detailing intellectual property and any public licensing and an associated analysis as to whether a compliant overarching open license can be applied

Data model

A data model presented graphically if possible, would help third parties understand the structure of the data. Such an artifact will have been produced as part of the model building exercise in any case.

Infrastructure and metadata

Spreadsheets, in this context, are inadequate in terms of data management. There are many more sophisticated methods on offer today and I suggest that the data mesh model from Jean-Georges Perrin is one such method.

Data meshes support metadata, provenience, revision history, fine-grained attribution, and fine-grain legal contexts — all missing from the current PRIMES databank. Please see:

- Perrin, Jean-Georges and Eric Broda (8 October 2024). *Implementing data mesh: design, build, and implement data contracts, data products, and data mesh*. Sebastopol, California, USA: O'Reilly Media. ISBN 978-1-09-815622-0.

Copyright, 96/9/EC database rights, and public licensing

Good practice transparent public policy includes the legal right for third parties to use and reuse the underpinning information for any purpose. That legal right needs to be transmitted through the use of public licensing that meets the Open Definition from the Open Knowledge Foundation.

Furthermore, the semantic standards in play need to be provided and suitably licensed. I suggest, using the SPDX system, "CC0-1.0 OR MIT" for this purpose. The data model mentioned earlier is the semantic standard most likely appropriate in this context.

Closure

This submission highlights the serious public interest deficiencies resulting through the use of closed frameworks and public but otherwise legally encumbered databanks. These deficiencies need to be rectified. The European Commission publics deserve better.

I should add that this submission was prepared at short-notice and that I would be happy to elaborate on the points raised.

My appreciation to the European Commission for being able to offer my views on the vitally important matter of future energy systems and analysis covering options for their rapid decarbonization.

Copyright (c) 2024 Robbie Morrison robbie.morrison@posteo.de
This work is licensed under a Creative Commons Attribution 4.0 International (CC-BY-4.0) License.