EUREC Input: Data Sharing as Foundation for the EU's Data Union Strategy

We welcome the European Commission's interest and heightened importance of data sharing and digitalization, especially with regard to renewable energy systems and particularly concerning the use of data in Artificial Intelligence (AI). In order to ensure coherence between policies, infrastructures, and legal instruments on data, clear and comprehensive data sharing policies must be the first step. An effective AI ecosystem can only be built using a large pool of trusted data in a common space. Relevant data should be publicly available, disaggregated (but anonymised for privacy concerns), and stored in accessible and scalable data formats.

Applying data-driven digital tools in the form of AI-driven software can optimise O&M strategies and renewable energy plant designs, ultimately driving down the cost of reaching 2030 targets[1]. These tools would:

- optimise intervals for plant repairs and pinpoint the cost-optimal repair
- detect anomalies in plant performance and advise on likely causes
- give insights into the optimal combinations of subcomponents in a plant and plant layout

Digital tools require a strong foundation of sharing operational data

With greater data sharing capabilities, large datasets could be exploited by AI technologies to optimise the performance and predictability of generation assets. Data can also be used to prepare improved models of the energy system and its components. For the AI-driven tools to perform well, high-quality[2] data from tens of GW of plants will be needed to train them. Such volumes of data are greater than all but the biggest wind or PV plant owners have access to, so the tool creator will most often need several data owners to share data with them. But data has value, so the owner must be offered something in return. The data will have to be operational, real-time data, such as electricity output, light levels, wind speeds, ambient temperature; and meta-data describing the plant. Meta-data is especially crucial and must be standardised to enable the uniform description of plant features and especially topologies in the physical, control, and ICT domains; meta-data is key to producing context-related information (e.g. for diagnostics) regarding the operation of a plant.

Operational data is valuable, but this value is often not realised. A use case we would like to see considered alongside those in the Digitalisation of Energy Action Plan is the use of operational data to optimise O&M strategies and plant designs, ultimately driving down the cost of reaching 2030 targets. With greater data sharing capabilities, larger quantities of data would be available to be exploited by AI (or non-AI) technologies to optimise the performance and predictability of generation assets.

How to incentivise sharing?

There are five clear ways in which data might be shared between a data owner and one or more recipients, depending on the sensitivity of the data. The table below explains all five, as well as a presumed price premium the data owner would demand for each case. Data naturally has value, and therefore also has a price. Industry representatives in a February 2022 workshop organised by EUREC and DG ENER noted that "Data owners would gladly share data if a price tag would be added to it."

[1] Most concretely, the Renewable Energy Directive target of 42.5% of final energy consumption from renewables by 2030.

[2] Data must be FAIR (Findable, Accessible, Interoperable and Reusable).

Case	Type of sharing	Premium
1	Operational plant data shared publicly; the plant and its components identified[3]	High
2	As Case 1, but data is anonymised just enough to make a specific plant unidentifiable[4]	Low
3	Operational plant data shared with a group of recipients but not publicly; precise identities of the recipients not known at time of winning the tender, only their nature; the plant and its components identified[5]	Medium
4	As 3 but the recipient(s) is/are named at the time of bidding. Recipients' ability to work with another provider of data is restricted[6]	Very low
5	As 4 but recipients' ability to work with other providers of data is unrestricted[7]	Very low

Table 1 Scenarios ('cases') for ways in which the owner of operational data from a renewable energy installation might share it with another party or parties. The right-hand column qualitatively describes the premium the owner might add to its bid compared to tender making no requirements for sharing, relative to the other scenarios.

Use non-price criteria to incentivise data sharing

Auctions with non-price criteria are simple yet effective methods of incorporating data-friendly initiatives. For example, the Net Zero Industry Act (NZIA) will see Member States use non-price criteria[8] that support innovation in auctions awarding public support to renewable energy installations (alongside criteria that support resilience, sustainability, and grid integration). Through non-price-criteria auctions, operators may state their price for the conditions laid down in the auction. Crucially, no company would be forced to participate in the auction, but auctions would indicate the market's appetite for data sharing. Therefore, data-sharing activities should qualify for meeting the innovation non-price criteria; the data sharing regime can be clearly specified according to the auction, where the less the data is anonymised and the more widely it is shared, the greater the points that can be won in the evaluation of the bid.

As NZIA is already in the implementation process, the next Governance Regulation review (expected in 2026/2027) could incorporate some of these ideas surrounding data-sharing criteria in auctions. Member States could also be obligated to report data sharing incentives/initiatives in their national strategies, including National Energy and Climate Plans (NECPs). In any case, the Commission should

[3] Case 1 corresponds to the model of the US Dept of Energy Solar Energy technologies Office Solar Data Bounty Prize (2023). Anyone may download the multi-gigabyte datasets of the participating plants.

[4] Case 2 allows anyone accessing the data to benchmark a specific plant against a plant having the same characteristics, albeit more approximately than cases 3-5 because the data is anonymised, making it impossible to know the full extent to which like is being compared with like.

[5] In case 3, the challenge is to define classes of recipient having a legitimate interest to access the data. Public research centres listed on a specific register or that have published, or presented, a recent paper in a relevant journal or conference could be eligible; companies might need to show they are developing plant monitoring software or hardware to qualify as recipients. The question of whether those companies are under European control is relevant

[6] For case 4, a very low premium would be needed because the relationship between the provider data and recipient will be very tight, and exclusive (at least for the recipient). The recipient will provide insights to the provider about their plant – that will be motivation enough for the provider to share. In return, the recipient will be able to use those insights for the improvement of its own tools.

[7] In case 5, the situation where the recipient is a government body has been tested in the US since 2019, when the DoE started funding the PV Fleet Performance Data Initiative. If a plant owner agrees to share detailed plant information with the Initiative, DoE scientists will compare the particular plant against the fleet and report the result to the owner. Private companies offering plant monitoring services are also examples.

[8] Refer to NZIA Article 26 and 28 for non-price criteria on innovation, alongside criteria that support resilience, sustainability and grid integration.

explore all avenues available use non-price criteria to encourage Member States to set operational data sharing criteria for some modest proportion of newly-installed renewable energy sources.

Use the Common European Energy Data Space framework

The Common European Energy Data Space (CEEDS), is a promising initiative that could help EU industry to rally behind a common data sharing architecture. CEEDS is a system for entities to share energy data in full confidence that the rules under which the sharing happens are understood and respected by the parties concerned, and lawful. CEEDS, if widely utilised, will make sharing even easier and more common.