**Annex 1: Data Anonymization Specification for PV Data**

**1. Introduction**

This document details the methods and protocols to be employed to ensure the photovoltaic (PV) data used in COPLASIMON is anonymized. This process is vital for safeguarding specific site details, such as location, while preserving the general utility of the data.

**2. Objective**

The primary objective of this data anonymization process is to prevent the identification of individual PV installation sites and other potentially identifiable data, ensuring data security and relevance for our project purposes.

**3. Data Anonymization Methods**

Several methods have been identified and, following the discussion with the external partners, they will be applied on the data:

**3.1 Location Obfuscation**: To protect the identification of specific PV installation sites, the precise location data will either be excluded or generalized to a lower resolution (e.g., city or district level rather than exact coordinates).

**3.2 De-identification**: All direct identifiers that might point to a particular installation or entity will be removed.

**3.3 Pseudonymization**: Original identifiers will be replaced with artificial ones, ensuring that backtracking to the original data is not feasible.

**3.4 Aggregation**: Data will be summarized in such a way that individual data points related to specific PV installations cannot be isolated.

**3.5 Data Masking:** Techniques like data shuffling or substitution will be used to obscure specific data within the dataset.

**3.6 Differential Privacy**: Random noise might be introduced to the data to protect the identification of particular PV installations while retaining the overall trends and patterns in the data.

**4. Approval on the Anonymization**

The data will be resubmitted to the partner for approval before any publication or distribution through COPLASIMON.

**5. Conclusion**

By adopting the methods described above, we ensure that the PV data used in COPLASIMON remains relevant for analysis without compromising the privacy and security of specific installation sites.