Pseudonymization Terminology for Policy Makers

Version 0.9.20, feedback to research@datenschutzzentrum.de

pseudonymization: (defined in Art. 4(5) GDPR)
A **manner of processing** in which directly identifying data elements (*additional information*) are kept separate and protected against unauthorized use in order to prevent the identification of data subjects during the processing of *pseudonymous data*.

data pseudonymization:
Data pseudonymization is a transformation of fully identified *personal data* that separates *pseudonymous data* and *identifying information*

pseudonymous data:
Pseudonymous data is *personal data* in which data subjects cannot be identified without the use of *additional information*; *identifying data* that results from *data pseudonymization* is one kind of *additional information*.

additional information: (defined in Art. 4(5) GDPR)
Additional information is any information suited to be combined (typically by linked) with *pseudonymized data* in order to identify (at least some) data subjects. One kind of *additional information* is the *identifying information* that results from *data pseudonymization*; other kinds of suitable *additional information* can exist and be held either by the controller or by external parties.

identifying information:
Identifying information is a kind of *additional information* that is the result of *data pseudonymization* and is kept separately and protected during *pseudonymization*. It permits to establish a one- or bi-directional relation between fully identifying data elements and the *pseudonyms* used in *pseudonymous data*.

pseudonymization reversal information: (used in Art. 44(3) EHDS)
Pseudonymization reversal information is bi-directional *identifying information* that permits to map from *pseudonyms* to fully identifying data elements.

pseudonymization reversal: (used in Art. 44(3) EHDS)
Pseudonymization reversal is the inverse of *data pseudonymization* that maps *pseudonymous data* plus *pseudonymization reversal information* to fully identified personal data.

reversible pseudonymization:
Reversible pseudonymization is *pseudonymization* in which the *pseudonymization reversal information* is kept available to enable a full or partial *pseudonymization reversal*.

irreversible pseudonymization:
Irreversible pseudonymization is *pseudonymization* in which the *pseudonymization reversal information* is not or no longer kept such that the controller is unable to perform a full or partial *pseudonymization reversal*.

pseudonym:
A *pseudonym* is a handle for data subjects used on both the *pseudonymous data* and the *identifying information*.

pseudonym scheme:
A *pseudonym scheme* is the manner in which *pseudonyms* are created during *data pseudonymization*.

pseudonym domain: The context in which a single *pseudonym scheme* is applied and consequently, each data subject is identified by a unique pseudonym that allows linking of data elements belonging to the same data subject.

2nd-level pseudonymization:
2nd-level pseudonymization is a transformation that replaces the *pseudonyms* in *pseudonymous data* with newly created ones. It uses a separate *pseudonym scheme* to create a distinct ("unlinkable") *pseudonym domain*.