

# Personal data and the Open Research Data Pilot

## How can OpenAIRE help?

Briefing paper for researchers,  
research administrators and  
project coordinators

### The EC Open Research Data Pilot

Open data is data that is free to use, reuse, and redistribute. The EC Open Research Data Pilot aims to make the research data generated by Horizon 2020 projects open. The Pilot applies primarily to the data (and metadata) needed to validate results in scientific publications, as well as other data specified in a data management plan (DMP).

Projects participating in the Pilot are required to deposit the research data in a research data repository and take measures to enable third parties to access, mine, exploit, reproduce and disseminate this research data.

However, the concept of the free use of research data within the Pilot may conflict with data protection rules if such data contain personal data.

### What is personal data?



“Personal data” means any information relating to a natural person who is either identified or who could be identifiable by that data (e.g., by reference to an identifier such as a name, an identification number, location data, online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that person).

Data protection rules always apply wherever personal data is processed. Processing here includes practically *any* operation in connection with personal data – including collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction.

## Dealing with Personal Data

### How to balance Open Access and Data Protection?

#### So what's the problem?

Research data – especially in fields like medicine, biotechnology and the social sciences – often contain personal data. This means that many datasets, in their raw form, cannot be made available on an Open Access basis as is required by the Open Research Data Pilot due to conflicts with rules on protection of personal data. Hence, incompatibility with data protection regulations is one of the major reasons for opting out of the Open Research Data Pilot. However, opting out of the Pilot is not the only way to prevent possible infringements. Firstly, because even if one particular dataset is unsuitable for sharing, the same project might produce other datasets which are. Secondly, because even if datasets contain personal data, they might still be able to be shared either through (1)

anonymisation, or in limited cases, (2) targeted sharing.

*Anonymised data are “data rendered anonymous in such a way that the data subject is no longer identifiable”*

(recital 26 of Directive 95/46/EC).

#### Anonymisation of personal data?

The best way to fulfil the requirements of the Open Research Data Pilot and data protection rules at the same time is to anonymise personal (research) data before making them openly available. Anonymised data are no longer personal data, which means that data protection rules are no longer applicable. Effective anonymisation prevents third parties from

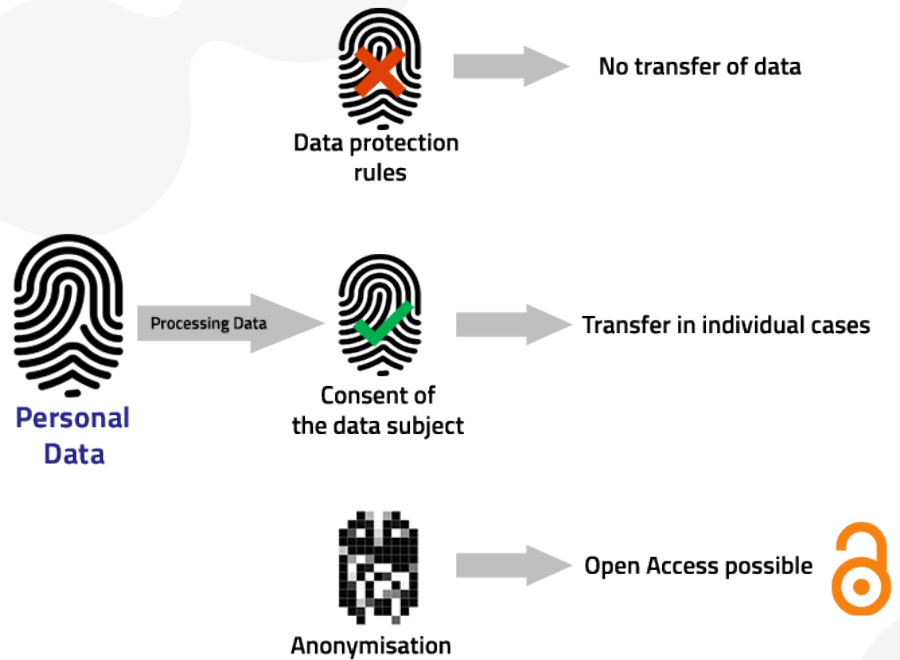
re-identifying individuals in anonymised datasets, i.e., associating a record to a natural person by using other sources of information. Moreover, anonymisation provides further **privacy guarantees** that prevent third parties from inferring that a person is associated with a certain property, e.g., a certain health condition, with a high probability or even to infer the participation of a person in a published dataset. When possible, anonymisation is the best solution to exclude data protection risks.

Coming soon:

**AMNESIA**, a service currently being developed by OpenAIRE, will allow data curators to anonymise their data.

## What if anonymisation is not possible?

Where complete anonymisation of the research data is not possible, another way to guarantee compliance with data protection rules is to obtain the consent of the data subject to use and exchange his data. According to Art. 2 (h) of Directive 95/46/EC and Art. 4 (11) of the upcoming data protection regulation, the consent of the data subject to processing his personal data must be freely given, specific, informed and unambiguous. To guarantee an informed consent, the purpose for processing must be defined. Since the purposes of the further use of data in an Open Access environment are unclear, it is not possible to legitimise the use of data in the Open Research Data Pilot by consent. However, it is at least possible to enable the transfer of personal data in an individual case.



Anonymisation of personal research data is the **only** effective solution to comply with **both** the data protection legislation and the requirements of the **Open Research Data Pilot**.



## OpenAIRE

### Training and Support for Open Data

#### How can OpenAIRE help?

OpenAIRE, in addition to the AMNESIA anonymisation tool currently under development, also provides a range of resources, FAQs, webinars and support pages, and has local representatives in all EU countries: the National Open Access Desks or NOADs.

Find resources and contact the NOADs via [www.openaire.eu](http://www.openaire.eu)

#### About OpenAIRE

OpenAIRE fosters the social and technical links that enable Open Science in Europe and beyond.

[www.openaire.eu](http://www.openaire.eu)

For more information, please contact:  
[info@openaire.eu](mailto:info@openaire.eu)

#### Other links

OpenAIRE Open Data Resources:  
[www.openaire.eu/opendatapilot](http://www.openaire.eu/opendatapilot)

EC's Guide on Open Access to Scientific Publications and Research Data in Horizon 2020:

[http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-pilot-guide\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf)

EC's Guidelines on Data Management in Horizon 2020:

[http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-data-mgt\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf)

DCC's DMPonline tool:

<https://dmponline.dcc.ac.uk>