

Using Administrative Data for Research

10
Legal and
Ethical issues
to consider



seriss
WWW.SERISS.EU

SYNERGIES FOR EUROPE'S
RESEARCH INFRASTRUCTURES
IN THE SOCIAL SCIENCES

This project has received funding from
the European Union's Horizon 2020
research and innovation programme
under grant agreement No 654221.



10 Legal and Ethical issues to consider

1 When do I need to think about the legality and ethics of administrative data linkage?

If you intend to link your research data to administrative data sources then it is optimal and often necessary to consider this at the very outset of your project. From a legal, ethical and technical perspective, the linkage to administrative data must be incorporated into the research design. From a legal perspective, linking to administrative data can often depend on the research project aims and scope aligning with the legally mandated objectives of the administrative data provider. To link with such administrative data, it can be helpful to ensure that the research question under investigation is in alignment with the aims and scope of the administrative data provider.

2 Who should I contact about the legal and ethical aspects of linkage?

Some administrative data providers have support for research, these organizations commonly have service desks who support administrative data linkage. This is the case for statistical offices whose general work is to support scientific research as well as statistical processing for government purposes. To further facilitate these service desks, many research councils have financed the development of support services for administrative data linking that are independent of administrative data providers.

3 When do I need consent for data linkage?

Under the General Data Protection Regulation (GDPR), personal data can be processed for several different reasons. For social research, there are two main legal bases on which personal data may be processed; because consent has been acquired or because it is in the public interest to do so. Research on administrative data does normally not require consent and data can be processed so long as specific protections are provided for. By contrast, survey data is predominantly processed based on the respondents' consent. When linking administrative data and survey data, you need to consider which legal basis to use.

4 What is required information to data subjects when processing personal data?

The GDPR lists the information that must be provided to a data subject in relation to the processing of their personal data where it is collected from the data subject (Article 13) or obtained from another source (Article 14). For an overview of the required information see the Article 29 working party guidelines on transparency, pages 35-40.

5 What is granular consent?

The best way to achieve informed consent for data sharing is to identify and explain the possible future uses of their data, and to offer the participants the option to consent on a granular level where relevant and possible. For administrative data linkage one approach could be to allow the participants the choice of what the data can be processed for. This could be based on explicit research project aims, the types of data that can be linked or the types of researchers who access the data. It is important to note that it is the responsibility of the data controller to oversee and monitor whatever granularity that is provided. It is therefore in the interest of the controller that such granularity is well defined, explicit and clear.

6 When do I process personal data?

The GDPR defines 'personal data' as any information that can be used to identify a person ('data subject'), directly or indirectly. When designing research with an administrative data linkage component, it is worth noting that identifiers for individuals and households that are used by the government have been included in the definition of personal data. This means that social insurance numbers, tax codes, driving license number and other identification numbers are all included under the definition of personal data. So long as an administrative linkage key exists, all data is considered personal data for processing purposes.

7 What can I do if I have consent from respondents to link data?

Administrative data providers are under no obligation to enable the linkage of social research data and administrative data even in instances where explicit consent is given by the data subject. This means that it is advisable to attain the cooperation of the relevant administrative data providers prior to data collection and the acquisition of consent to ensure that it would permit linkage on the legal basis under which the administrative data provider processes data.

8 When is administrative data anonymous?

Administrative data is only anonymized when the linkage key to all personal data is destroyed and reasonable precautions have been undertaken to prevent data disclosure. Certain administrative data can be redacted or masked in such a way as to mitigate disclosure risks, such as the collapsing of income data into percentiles or the use of established masking techniques. Specific recommendations on pseudonymization and general anonymization procedures are provided by the Article 29 working party.

9 How can I document my workflow in a secure environment?

Administrative data rarely comes with publicly available metadata. When engaging in administrative data linkage it is therefore important to develop a coherent workflow documentation plan at the outset before access is granted. Even if the researcher doesn't use such a documentation plan in their regular work, they are especially important in administrative data linkage. Furthermore, it is good practice to document the access procedure and seek clarification from the data controller about any plans for sustainability of the data access process that exist. This will facilitate replication of any analysis conducted.

10 How can I archive administrative data?

Administrative data can be rarely archived in a way that makes it accessible as administrative data providers are generally required to maintain full control over their data. Nevertheless, archiving the data at a data archive without the provision of access to other researchers is still exceptionally valuable. Even the storage of only the available metadata still satisfy the FAIR principles, especially if they clearly document and update the procedure and contact points for gaining access to the data. Here it is important to note the difference between FAIR and open data. It should not be presumed that the administrative data provider is an adequate long-term archive.



seriss

SYNERGIES FOR EUROPE'S
RESEARCH INFRASTRUCTURES
IN THE SOCIAL SCIENCES

Legal disclaimer: This does not constitute or should not be construed as legal advice and/or guidance.