EDPB Guidelines 1/2026 on processing of personal data for scientific research purposes

-

Feedback from Fondazione Bruno Kessler (FBK) and Università di Trento (Unitn)

We welcome the draft guidelines regarding the application of the General Data Protection Regulation (GDPR) within the scientific research sector. Given the exponential growth of data-intensive research, there is a pressing need for these guidelines to ensure that innovation and the GDPR are going hand-by-hand. While the guidelines provide many clarifications, we believe the following areas could require some more context to ensure they are functional applicable in the research field.

1. Scientific research purposes

We really appreciate the clarification offered by the draft guidelines on what can be considered scientific research. We specifically appreciate the mentioning of six 'key-indicative factors' to assess whether a particular research activity can be qualified as a scientific activity. It would be, however, useful if it could be clarified (in accordance with example 2 at p.11) if research organizations may be considered to carry out scientific research in partnership with private entities both when (i) participating to the same research project (e.g. within a consortium) and (ii) when carrying out R&D activities under a commercial contract that would benefit the scientific community (e.g. with publications or open access results), considering that in both situations the project maintains scientific features, such as methodological autonomy, qualified researchers, verifiability, ethical or independent oversight, and contribution to scientific knowledge.

It would also be useful to include further examples concerning AI and data-intensive research, in particular to distinguish between processing operations carried out for genuine scientific research - such as training, testing, validation, benchmarking, bias assessment or model evaluation - and processing operations mainly aimed at product development, internal optimisation or commercial analytics.

On the point of ancillary processing operations (section 2.3) which can be linked to the research purpose, we were wondering whether this also includes ancillary processing operations conducted by a different (independent) data controller. For example, if a company pseudonymizes a dataset before sharing this dataset with the research institute, does this operation concern an ancillary processing operation, hence the purpose of this operation of the company can be considered a "research" purpose?

2. Purpose limitation

The first comment relates to the question of the primary and secondary purpose of processing (section 3.1.1). In the draft guidelines it is written that if a controller collects personal data to process it for a specific scientific research purpose, then the research purpose is considered to be the primary purpose of processing. If a controller subsequently wants to process that personal data for another scientific research purpose that is not covered by the primary purpose, such processing is considered to be the further processing for another purpose. We see a tension between the principle of purpose specification and research projects. Should every individual research project be considered a specific purpose and the subsequent use of data from that project in a new project a further processing? Or can research, for example research in the field of AI and early detection of a particular type of cancer (i.e. breast cancer), be considered specific enough, meaning that all the projects in that field, where the same personal data will be processed, are considered the primary use of personal data?

With respect to the further processing when providing personal data to another controller for scientific research purposes (para 3.1.2), it is not entirely clear from the text how to interpret and apply this to the situation where there is a consortium of different research partners (e.g. in Horizon Europe projects) that act as joint controllers. It would be useful to distinguish between the provision of data to another controller for a new autonomous purpose and processing carried out by different partners within a jointly defined research purpose (i.e. where the common purpose is already defined in the grant agreement, research protocol, data management plan and/or consortium agreement). If for example, research partner A (a hospital) collects health data directly from the patients with the purpose of offering healthcare and research partner B (a university) processes them for their common goal of conducting research in a specific field, should we then consider the processing of partner B as a further processing knowing that they are both part of the same research project?

As a last point relating to the further processing, we would like to ask whether it is possible to discuss the distinction between primary and secondary use in light of other existing (horizontal) legal frameworks such as the European Health Data Space. Potential discrepancies between the GDPR's "further processing" and the EHDS's "secondary use" could potentially lead to fragmented compliance across the EU. The same clarification would be useful not only in relation to the EHDS, but also with regard to other EU data frameworks, such as the Data Governance Act and, where relevant, the Data Act. A short table or examples comparing GDPR "further processing", EHDS "secondary use", and data access or data sharing mechanisms under other EU instruments would help avoid fragmented compliance.

3. Consent

With respect to consent, the draft guidelines propose a distinction between broad consent and dynamic consent. With respect to broad consent, the draft guidelines state that the purpose

could be delimited to a certain field of research, for example medical research in the field of oncology (point 45). Does this imply, see the comments above, that in general this could be considered a valid “purpose”, meaning that the multiple use of certain personal data in different research projects falling in a certain research field could be considered the “primary” use of personal data? With specific application to research data infrastructures (section 2.2; par. 13), which refer to section 4.1.2.1 on broad consent, how much broader could that consent be?
It would be highly beneficial if the Guidelines could provide practical criteria to help determine the appropriate level of granularity when defining a “certain area of scientific research”. This interpretive support would be particularly valuable for research data infrastructures, biobanks, clinical imaging repositories, health data lakes and AI-based medical research infrastructures, where data may be used for future projects. For instance, if a patient undergoes an MRI or a blood panel for diagnostic purposes and gives consent to have the data included in a research data infrastructure, that exam could be used for a wide variety of research projects and domain branches in the field of medicine. Is it enough that the access criteria to the infrastructure are fulfilled? If so, should they be communicated to the patient when giving consent and what if they change over time due to new legal requirements or practices?

Further clarification would be welcome on how changes to access criteria, governance rules or eligible research areas should be handled over time. In particular, the Guidelines could distinguish between changes that merely require updated information to data subjects (section 5.5) and changes that require renewed or separate consent.

With respect to the category “dynamic” consent, it is not entirely clear what exactly is the dynamic character of this consent. By reading the explanation (section 4.1.2.2) this seems the “normal” process for asking consent, or renewing consent when the new data processing activity does not fall within the scope of the initial consent. It would be appreciated if the guidelines could clarify if "dynamic" refers to a specific legal modularity that creates certain flexibilities for the data controller and/or data subjects.

The Guidelines could also provide examples on the practical consequences of withdrawal of consent in scientific research, especially in relation to pseudonymised datasets, archived research data, trained AI models, peer review and reproducibility. This would help clarify when continued retention or use of certain data or derived outputs may remain justified.

Furthermore, the Guidelines (e.g. in example 8) could clarify how the information requirements relating to additional processing purposes are to be addressed where the data subject is deceased or can no longer be contacted. While we acknowledge that, in such circumstances, Member State law rather than the GDPR may apply, this remains a complex issue that would merit further clarification, particularly in cases involving university hospitals operating under different national legal frameworks.

4. Storage limitation

The Guidelines should provide more practical examples on post-project retention. While they recognise that data may be retained for verification, peer review and future specific research purposes, it remains unclear how data controllers should apply these principles in practice, especially in multi-project datasets and research infrastructures. Clarification would be useful on whether retention can be justified by scientific integrity requirements, such as reproducibility, audit, validation of results, grant checks or ethical oversight, and on whether a staged retention model based on increasing access restrictions and pseudonymisation would be acceptable.

5. Public interest

Pursuant to article 6(1)(e) and article 6(3) GDPR, processing of personal data in the public interest must be based on a legal basis laid down by Union or Member States law to which a data controller is subject. It is still a bit unclear how explicit the Union or Member States law should mention that the processing of personal data for scientific research purposes is allowed. The Guidelines could provide a practical interpretative matrix on whether Union or Member State law is sufficiently clear, precise and foreseeable to support processing for scientific research in the public interest. In particular, it would be useful to clarify whether a general institutional mission, public funding, sectoral legislation or EU policy objectives may be sufficient, or whether express authorisation to process personal data for scientific research purposes is required.

Below we have listed a few examples to understand better which activities could be considered as a public interest.

If a research project is funded by the European Union, for example an EU Horizon project, and this project develops an AI system that can find disinformation in social media platforms. Is that research project a public interest considering that tackling online disinformation is part of the strategy of the European Commission to shape Europe's digital future and mentioned in several documents (i.e. the Code of Practice on Disinformation) and in the Digital Service Act?

Another example is the Section 12 (10-14) of Italian Law Decree no. 179 of 18 October 2012 (which is mentioned in draft Guidelines). This law focuses on the establishment of disease-specific registries, data governance and the integration of these records into the National Health Service. There is no mention of the need to process personal data nor the processing of these personal data for scientific purposes. You can only infer from the activity that personal data will be processed. However, a link to processing this data for scientific purposes is missing. Is this data processing in the public interest?

A last example, which according to our reading, is a good example of a MS law stating that certain processing activities are in the public interest is Italian Law Degree no. 132 of 23 September 2025. Article 8 of this law states that "*The processing of data, including personal*

data, carried out by public and private non-profit entities, …, for scientific research and experimentation in the development of artificial intelligence systems for the prevention, diagnosis and treatment of diseases, the development of drugs, therapies and rehabilitation technologies, the manufacture of medical devices, including prostheses and interfaces between the body and instruments supporting the patient's condition, public health, personal safety, health and safety, as well as the study of physiology, biomechanics and human biology, including in non-health contexts, insofar as they are necessary for the creation and use of databases and basic models, are declared to be of significant public interest in implementation of Articles 32 and 33 of the Constitution and in compliance with the provisions of Article 9(2)(g) of Regulation (EU) 2016/679."

6. Processing special categories of personal data for scientific research purposes

The Guidelines should clarify that explicit consent under Article 9(2)(a) GDPR should not be treated as the default derogation for processing special categories of data in scientific research. Where Union or Member State law implements Article 9(2)(j) GDPR and provides suitable safeguards under Article 89 GDPR, data controllers should assess whether Article 9(2)(j) is the more appropriate basis. This would also help avoid confusion in privacy notices between consent to participate in research, consent as a GDPR legal basis or derogation, and the safeguards required by research-specific legal frameworks. Once, in line with Article 9(2)(j), a project applies national rules it might be possible that the national rules in any case require consent as the additional safeguard (see the examples quoted in footnote n. 115).

7. Obligation to inform

The Guidelines could mention the importance and potentiality of legal design of privacy policy with icons and signs able to clarify and render accessible information to the data subjects. Beyond the visual presentation, optimizing the wording and managing the length are equally critical, especially in a sensitive field like research.

8. Joint controllers

In practice, many researchers' institutes prefer to declare themselves autonomous-independent data controller to avoid the joint controllership agreement. This, even in the case of the same research protocol/project proposal and methodology (see example 25). They frequently rely on the fact that they transmit only pseudonymised data to the principal investigator and databases are not shared. The identification means are retained only by the data originators/first controller. The Guidelines could clarify whether this choice is compatible with the interpretation of the GDPR and, according to the answer, provide at least one example that better shows this context.

9. Appropriate safeguards (Article 89(1) GDPR)

Finally, the Guidelines would benefit from more practical examples of appropriate safeguards in complex research environments, such as Horizon Europe consortia, federated infrastructures,

health data spaces and AI research platforms. Examples could include data access committees, trusted research environments, access logs, audit trails, data access agreements, secure export of results, data trustees and independent oversight.
It would also be useful to clarify when privacy-enhancing technologies such as federated learning, synthetic data, differential privacy, secure multiparty computation and secure processing environments may qualify as appropriate safeguards under Article 89 GDPR, and how residual re-identification risks should be assessed.

Sincerely,
The working group

Fondazione Bruno Kessler

- Dr. Anne Elisabeth Beumer
- Dr. Giulia Olivato
- Ms. Anna Benedetti
- Ms. Valentina Campana

Università di Trento

- Prof. Paolo Guarda
- Dr. Giorgia Bincoletto