



European Research Council  
Executive Agency

Established by the European Commission

## RECORD OF PERSONAL DATA PROCESSING

Art. 31 of the REGULATION (EU) 2018/1725 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 23 October 2018 on the protection of natural persons with regard to the processing of personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data, and repealing Regulation (EC) No 45/2001 and Decision No 1247/2002/EC (henceforth the "Data protection regulation")

Record n°

DPO-62-2025

In accordance with Article 31 of the data protection regulation, individuals whose personal data are processed by the Executive Agency in any context whatsoever are to be protected with regard to the processing of personal data and the Executive Agency has to keep records of their processing operations.

This record covers two aspects:

1. Mandatory records under Art 31 of the data protection regulation (recommendation: make the header and part 1 publicly available)
2. Compliance check and risk screening (initial; part 2 is internal only to the Agency, not published)

The ground for the record is (tick the relevant one):

- Regularization of a data processing operation already carried out
- Record of a new data processing operation prior to its implementation
- Change of a data processing operation.

(This part may be public)  
**Classification of ERC projects through LLMs**

1	Last update of this record if applicable	N/A
2	Short description of the processing	The ERCEA intends to process the full proposal text of the funded ERC projects for the purpose of the Mapping Frontier Research (MFR) classification of ERC projects by testing and potentially operationalizing the use of Large Language Models (LLMs).

(This part may be public)  
**Part 1 - Article 31 Record**

3	Function and contact details of the controller	ERCEA Director or Head of Scientific Management Department. (Dpt B) by delegation: <a href="mailto:ERC-B-WORKSHOP@ec.europa.eu">ERC-B-WORKSHOP@ec.europa.eu</a>
---	--	---

4	<b>Contact details of the Data Protection Officer (DPO)</b>	<a href="mailto:erc-data-protection@ec.europa.eu">erc-data-protection@ec.europa.eu</a>
5	<b>Name and contact details of joint controller (where applicable)</b>	N/A
6	<b>Name and contact details of processor (where applicable)</b>	GPT@JRC led by Dir. T supported by Dir. S. Contact: <a href="mailto:JRC-DIR-T-GPT@ec.europa.eu">JRC-DIR-T-GPT@ec.europa.eu</a> GPT@EC led by DIGIT.B. Contact: <a href="mailto:EC-DIGIT-GPT@ec.europa.eu">EC-DIGIT-GPT@ec.europa.eu</a>
7	<b>Purpose of the processing</b>	The purpose of the processing of funded projects' full proposal data (scientific section A of part B1 and part B2) is to test the feasibility of – and potentially operationalize - the use of LLMs (generative AI) for the classification of ERC projects into the MFR dimensions Disciplines (175 terms), Topics (584 terms) and Methods (162 terms) of the MFR taxonomy designed by the ERC Scientific Council. The funded projects' proposal data (scientific section A) is put in an LLM prompt, processed by the LLM and the LLM output consist only of Disciplines, Topics or Methods terms for each project with eventually some general text explaining the choice of these terms. The LLM output doesn't not contain any personal data. No data is kept by any of the LLM providers nor used for further training of the LLM by the providers.
8	<b>Description of the categories of data subjects</b>	<input type="checkbox"/> EA staff (Contractual and temporary staff in active position) <input type="checkbox"/> Visitors to the EA <input type="checkbox"/> Contractors providing goods or services <input type="checkbox"/> Applicants <input type="checkbox"/> Relatives of the data subject <input type="checkbox"/> Complainants, correspondents and enquirers <input type="checkbox"/> Witnesses <input checked="" type="checkbox"/> Beneficiaries Principal Investigators (PIs) since FP7 (technical debt if decision is made to do MFR classification on FP7) and current PIs concerned with the yearly MFR classification exercise <input type="checkbox"/> External experts <input type="checkbox"/> Contractors <input type="checkbox"/> Other, please specify:

9

**Description of personal data categories**

Indicate **all** the categories of personal data processed and specify which personal data are being processed for each category (between brackets under/next to each category):

Other : please specify : **Funded projects' proposal data.**  
**Only the scientific part of the proposal data will serve to classify the projects.**

*Categories of personal data:*

- in the form of personal identification numbers
- concerning the physical characteristics of persons as well as the image, voice or fingerprints
- concerning the data subject's private sphere
- concerning pay, allowances and bank accounts
- concerning recruitment and contracts
- concerning the data subject's family
- concerning the data subject's career (academic CV in project proposal)
- concerning leave and absences
- concerning missions and journeys
- concerning social security and pensions
- concerning expenses and medical benefits
- concerning telephone numbers and communications
- concerning names and addresses (including email addresses)
- Other :please specify

*Categories of personal data processing likely to present specific risks:*

- data relating to suspected offences, offences, criminal convictions or security measures
- data being used to evaluate personal aspects of the data subject (ability, efficiency, conduct)

*Categories of personal data whose processing is prohibited, with exceptions (art. 10 new Regulation):*

- revealing racial or ethnic origin
- revealing political opinions
- revealing religious or philosophical beliefs
- revealing trade-union membership
- concerning health

10	<b>Retention time (time limit for keeping the personal data)</b>	<input type="checkbox"/> genetic data, biometric data for the purpose of uniquely identifying a natural person <input type="checkbox"/> concerning sex life or sexual orientation <i>Specify any additional data or explanatory information on the data being processed, if any:</i>  The proposal data in raw text format of year N is kept until the MFR classification of year N is completed (no classification errors detected) and may be deleted once the exercise of year N + 1 is launched unless it is needed to make some comparison between 2 years. After 2 years, the proposal data in raw text format is deleted.
11	<b>Recipients of the data</b>	The MFR classification data generated by the LLM (disciplines, topics, methods) is only used by ERCEA staff (Dpt B scientific officers and B0 F2P staff) and made available through MFR tool (D1).
12	<b>Are there any transfers of personal data to third countries or international organisations? If so, to which ones and with which safeguards?</b>	N/A
13	<b><u>General</u> description of the technical and organisational security measures</b>	The general technical and security measures for protecting funded projects' proposal data is assured by the DIGIT central IT system security policy (SEP system). In addition, no data is kept by any of the LLM providers (data processors cf. section 6 above) nor used for further training of the LLM by the providers (cf. The public EC DPO register <a href="https://ec.europa.eu/dpo-register/detail/DPR-EC-20911.3">https://ec.europa.eu/dpo-register/detail/DPR-EC-20911.3</a> for GPT@JRC and "Data, privacy, and security for Azure OpenAI service" <a href="https://learn.microsoft.com/en-us/legal/cognitive-services/openai/data-privacy?tabs=azure-portal">https://learn.microsoft.com/en-us/legal/cognitive-services/openai/data-privacy?tabs=azure-portal</a> )
14	<b>Information to data subjects/Specific Privacy Statement (SPS)</b>	A data protection notice is published on the ERC website