



EDPS Opinion on the ERCEA note concerning the application of provisions on retention time for scientific, historical research and statistical purposes (Case 2021-0639)

1. INTRODUCTION

This Opinion relates to a note concerning the application of provisions on retention time for scientific, historical research and statistical purposes (the ‘Note’) submitted by the European Research Council Executive Agency (‘ERCEA’), complemented with nine annexes¹, on 17 June 2021.

The EDPS issues this Opinion in accordance with Articles 57(1)(g) and 58(3)(c) of Regulation (EU) 2018/1725², (‘the Regulation’).

2. BACKGROUND INFORMATION

In the context of its management of calls for proposals and grants, ERCEA processes significant amounts of personal data of EU civil servants and independent experts involved in the evaluation of research proposals and the assessment of the output of the funded projects.³

The EDPS issued prior checking opinions on data processing operations in the context of calls for proposals and grants management in 2011⁴ and regarding expert management in 2014⁵.

In the past years, ERCEA has been analysing the needs of storing the personal data collected in the framework of their activities and further processing them, under appropriate safeguards, for historical, statistical or scientific research purposes, in accordance with Articles 4(1)(b) and (c) and 13 of the Regulation.⁶

¹ - Legal framework applicable to the ERC/ERCEA and their activities (Annex 1);
- Legal basis applicable to the ERCEA processing personal data for the purpose of treating allegations of scientific misconduct (Annex 2);
- Detailed description and analysis of each purpose for processing and further processing of personal data (Annex 3);
- Proposed retention periods by purpose and type of data (Annex 4);
- IT measures and IT Draft protocol on transfer of data from the active database to the archive database (Annex 5);
- Draft data protection notice (Annex 6);
- Retention periods of other European Funding Agencies (Annex 7);
- ERC Board position regarding the retention and storage of data for scientific, historical research and statistical analysis (Annex 8);
- ERC Board Minutes 21 May 2019 (Annex 9).

² Regulation (EU) 2018/1725 of the European Parliament and 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, OJ, L 295, 21.11.2018, pp. 39-98.

³ Note, p. 2.

⁴ Case 2011-0845:
https://edps.europa.eu/data-protection/our-work/publications/opinions-prior-check/evaluation-and-grants-management-ercea_en

⁵ Case 2013-0575:
https://edps.europa.eu/data-protection/our-work/publications/opinions-prior-check/experts-selection-and-management-ercea_en

⁶ Note, p. 3.

The Note seeks for the EDPS formal opinion on ERCEA's practice and approach proposed for the **further processing of data for historical research, scientific research and statistical purposes, after the expiry of the operational retention time**, in accordance with Article 13 of the Regulation. According to ERCEA, the purpose of the Note is 'to present to the EDPS the well-thought reasons of the European Research Council Executive Agency (ERCEA) for retaining and processing personal data for scientific, historical research and statistical purposes and for identifying accordingly the retention periods it deems necessary to be able to fulfil its tasks in the interest of the public'⁷.

The Note also asks the opinion of the EDPS on the retention period of the data that needs to be processed for the **purpose of detection of plagiarism and any other scientific misconduct**, also beyond the operational retention time.⁸

On 22 January 2021, the EDPS provided informal comments on a first version of the Note submitted by ERCEA.

On 17 June 2021, ERCEA submitted the formal consultation on the Note to the EDPS.

3. LEGAL ANALYSIS AND RECOMMENDATIONS

3.1. Scope of the Opinion

The EDPS notes that most of its informal comments have been addressed in the Note and its Annexes.

This Opinion focusses on the aspects that raises issue of compliance with the Regulation or otherwise merit further analysis as regards:

- further processing of data collected for the purpose of detection of plagiarism and any other scientific misconduct, beyond the retention time for call evaluation and grant management;
- further processing of the same personal data after the expiry of the operational retention time for call evaluation and grant management, in order to pursue scientific, statistical and historical research purposes.

3.2. Further processing for the purpose of detecting plagiarism and any other scientific misconduct

3.2.1. Further processing

According to Article 6 of the Regulation, where the processing for a purpose other than that for which the personal data have been collected is not based on the data subject's consent or on Union law, a number of factors need to be taken into account in order to determine whether this further purpose is compatible or not. Namely, the controller shall, in order to ascertain whether processing for another purpose is compatible with the purpose for which the personal data are initially collected, should take into account, inter alia:

- (a) any link between the purposes for which the personal data have been collected and the purposes of the intended further processing;

⁷ Note, p. 2.

⁸ Note, p. 3 and pp. 9-10.

- (b) the context in which the personal data have been collected, in particular regarding the relationship between data subjects and the controller;
- (c) the nature of the personal data, in particular whether special categories of personal data are processed, pursuant to Article 10, or whether personal data related to criminal convictions and offences are processed, pursuant to Article 11;
- (d) the possible consequences of the intended further processing for data subjects;
- (e) the existence of appropriate safeguards, which may include encryption or pseudonymisation.

According to the Opinion 203 of the Working Party 29 (WP29)⁹, the nature of the assessment to be carried out by the data controller is decisive. In very brief terms, it can take two different forms. The compatibility test could be formal or substantive:

- A formal assessment will compare the purposes that were initially provided, usually in writing, by the data controller with any further uses to find out whether these uses were covered (explicitly or implicitly).
- A substantive assessment will go beyond formal statements to identify both the new and the original purpose, taking into account the way they are (or should be) understood, depending on the context and other factors.

Annex 3 to the Note communicated by ERCEA provides a detailed analysis of each purpose, including an assessment of the impact on the data subjects. According to ERCEA, the overall aim is to avoid that scientific misconduct jeopardises the value of science and in particular the reputation of the scientists in the scientific community, as well as of the bodies funding or hosting these scientists.

In addition, Annex 2 provides the legal basis applicable to the ERCEA for processing personal data for the purpose of treating allegations of scientific misconduct.

Based on the analysis conducted by ERCEA, the purpose of this specific processing seems **compatible** with the initial purpose (proposal evaluation and grant management), as it aims at eliminating potential forms of scientific misconduct and therefore, ensures the quality of the proposals financed by the ERC.

3.2.2. Appropriate safeguards

Article 6(e) of the Regulation states that the controller shall, in order to ascertain whether processing for another purpose is compatible with the purpose for which the personal data are initially collected, also take into account, the existence of **appropriate safeguards**.

According to Annex 3 to the Note communicated by ERCEA, there are appropriate specific access rights and controls following the “need to know” principle for data related to detection of plagiarism. Any electronic information is stored in a database that resides on the servers of the ERCEA and the European Commission, the operations of which abide by the European Commission’s security decisions and provisions established by the Directorate of Security for this kind of servers and services.

Data on scientific misconduct cases are recorded in the ERCEA ARES archive system with confidentiality markings, depicted by the case number only and with restricted access to ERCEA staff members dealing with the allegations. Furthermore, according to ERCEA, paper files are stored in a safe box with restricted access rights.

⁹ Article 29 Working Party, Opinion 03/2013 on purpose limitation, p. 21, available at https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf

3.2.3. Retention period

According to Section IV of the Note, ERCEA plans to apply a 15 years retention period, beyond the initial processing, for the purpose of detection of plagiarism and other scientific misconduct. In addition, Annex 4 provides a table with the different proposed retention periods by purpose and type of data and the justification for such extended retention time.

According to ERCEA¹⁰, a period of 15 years for processing personal data for all (funded and unfunded) applications is considered necessary in order for ERCEA to be able to verify the originality of the documents (primarily PhD certificates) as part of the eligibility conditions for applying to Starting and Consolidator grants. Another reason to retain data (proposals) for 15 years is to detect all cases of plagiarism.

In view of these elements, The EDPS finds the proposed retention period justified.

The EDPS also welcomes the fact that ERCEA has included the information on this further purpose in annex “ERC data retention and rights of data subject” of the draft Data Protection Notice (Annex 6). The EDPS reminds that this information should also be included in the relevant **record** of processing activities, in accordance with Article 31 of the Regulation (**Recommendation 1**).

3.3. Further processing for the purposes of scientific, historical research and statistical analysis

3.3.1. Appropriate safeguards

According to the purpose limitation principle, under Article 4(1)(b) of the Regulation, personal data shall be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes. Further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 13, not be considered to be incompatible with the initial purposes⁷.

The presumption is not a general authorisation to further process data in all cases for historical, statistical or scientific purposes. Each case must be considered on its own merits and circumstances¹¹.

ERCEA provides a number of detailed arguments in support of the various purposes for further processing on the Note and on Annex 3.

Article 13 of the Regulation states that processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, shall be subject to appropriate safeguards, in accordance with this Regulation, for the rights and freedoms of the data subject. Those safeguards shall ensure that technical and organisational measures are in place in particular in order to ensure respect for the principle of data minimisation. Appropriate safeguards could include data minimisation through pseudonymisation or (unless it would impair the research) anonymisation¹².

¹⁰ Annex 3, page 6

¹¹ EDPS Preliminary Opinion on data protection and scientific research, p. 22, available at: https://edps.europa.eu/sites/edp/files/publication/20-01-06_opinion_research_en.pdf.

¹² Ibid

Data minimisation and pseudonymisation

The EDPS welcomes the fact that the ERCEA will apply data minimisation for statistical purpose¹³.

As regards historical and scientific research, ERCEA underlines that creativity is a fundamental pillar of the research and that the nature of the future exploitation of the data by scientists is therefore highly unpredictable.¹⁴ Therefore, according to ERCEA, uncertainty about such requests recommends storing the data in their integrity without anonymisation, any selection or filtering.¹⁵ ERCEA also mentions¹⁶ that detaching the person (e.g. Principal Investigator, expert) from the rest of the information (e.g. proposal, expert opinion) would impair the ERCEA's capacity to fulfil the historical and scientific research purposes and would make the rest of the data meaningless. The EDPS understands the need to keep the data in their integrity and to identify the research subjects in that context but recommends that ERCEA explore the possibility to **apply pseudonymisation (Recommendation 2)**.

Additional technical and organisational measures

Processing for 'historical' and 'scientific' purposes can also have specific characteristics that often call for safeguards beyond anonymisation or pseudonymisation, including appropriate security measures and restrictions on access.

Annex 5 describes the archive system of the ERCEA and the various roles of users that can interact with it. The system relies on a set of 3 Oracle databases: one for storing the archive data (ERC_SECURED); one for auditing purposes (ERC_SECURED_LOG) and one for the Archive Manager to manage the access of users (ERC_SECURED_USERS). The Archive Manager acts as an interface between the Archive Owner (the data controller) and the software development team and is also responsible for granting/revoking access to the archive database to Authorised archive users (e.g., researchers and analysts).

Access to the data

Once the Archive Owner approves a request for access, the Archive Manager will be requested to implement this access. First, the Archive Manager provides a password for the authorized user to log into the database and later grants access via the authorised business roles to the authorised user and confirms to the user that they are able to view the appropriate data. Once the user has finished with the analysis, they will inform the Archive Manager who will revoke the data access and then change the password to the user's account.

In the perspective of the EDPS, the aforementioned procedure seems complex to manage, especially if the ERCEA happens to receive multiple simultaneous requests in a short timeframe. The Archive Manager (or his/her backup) has to ensure that every new user is granted the right permissions to access the authorized data (not less and not more) while, at the same time, making sure that the provided credentials are revoked as soon as they are not needed. This puts a heavy burden on the Archive Manager, which can lead to errors, ultimately resulting in excessive permissions and unauthorized access to personal data.

¹³ See Note, p. 8 and Annex 3, point e)

¹⁴ Note, page 8

¹⁵ Note, page 8

¹⁶ Note, page 8

In case ERCEA receives numerous requests the EDPS recommends that ERCEA **re-evaluate the proposed archiving architecture**, which currently relies heavily in the role of the Archive Manager to ensure the principle of data minimisation. At a minimum, the role of the Archive Manager should be undertaken by a team of experts, and not by a sole individual. This would allow for the revision of the permissions granted by different elements the team, reducing the likeability of errors (also known as the four eyes principle).
(Recommendation 3)

Prevention of copying

The archive system allows for users to make “temporary local copies”. The mechanism for prevention of (permanent) copies relies solely in an obligation on the signed request, where the user commits to the deletion of the data after the agreed duration of the request. Sample checks would be implemented to verify the compliance.

Annex 5 does not explicitly mention what information could be copied locally, but it is fair to assume that the user can make a copy of the database he/she has permission to access. The EDPS considers that the formal obligation is an insufficient mechanism, which does not properly guarantee the confidentiality of personal data. Furthermore, the possibility to make unabridged snapshots of databases outside of the environment of the data controller contradicts the principle of data minimization and presents serious dangers to the confidentiality of data, since the system of the user might not provide sufficient safety mechanisms against attacks or data theft.

The EDPS recommends that ERCEA **change the archive system** in order to provide exports of data, instead of copies of the databases. Before the export, each user should be prompted to indicate the specific records he/she wishes to export. The number of exportable records should be limited and any export outside the defined boundary should require prior authorization from the ERCEA **(Recommendation 4)**.

Finally, it is not clear how the ERCEA plans to implement the “**sample checks**” for **compliance verification** in the user side. This should be clarified in the Annex **(Recommendation 5)**.

3.3.2. Retention period

According to the storage limitation principle, under Article 4(1)(e) of the Regulation, personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 13 subject to implementation of the appropriate technical and organisational measures required by this Regulation in order to safeguard the rights and freedoms of the data subject.

According to Section IV of the Note, ERCEA plans to apply an indefinite storage period for the purpose of **historical research**. In Annex 3, ERCEA highlights that data shall be retained for an indefinite period to be able to provide future generations with the possibility to study the history of the ERCEA and of its applicants¹⁷.

Similarly, according to Section IV of the Note, ERCEA plans to apply a 25 years storage period for **statistical and scientific research** purposes. In respect of **statistical purposes**,

¹⁷ page 9

ERCEA highlights¹⁸ that data analysis is crucial to guide and improve the ERC program operations. Data helps the ERC and ERCEA to choose among competing strategies, redefine priorities, adjust to new trends, rapidly react to changes with the overall aim to achieve the policy objectives in the most effective and efficient way. ERCEA produced data are also used to evaluate outcomes, the effectiveness of the program, the impact it makes, which are part of the general accountability to the public.

Regarding **scientific research**, ERCEA states¹⁹ that since research can be a long-term endeavour and the scientific method tends to expand on past discoveries, preserving the ERCEA-held scientific evidence ensures that today's traces remain available for other scholars to consult and build on. Additionally, this could be relevant for protecting intellectual property and managing scientific information.

Given the above, data shall be retained for 25 years to be able to carry out relevant processing operations for these purposes, which are necessary for the ERCEA to be able to fulfil its mission²⁰.

The EDPS finds the proposed retention period justified provided that the appropriate safeguards are implemented taking into account its above recommendations 2 to 5.

Finally, these retention periods are also reflected on in the annex “ERC data retention and rights of data subject” of the draft Data Protection Notice (Annex 6). The EDPS reminds that this information should also be included in the relevant record of processing activities, in accordance with Article 31 of the Regulation (see **Recommendation 1**).

* *
*

4. CONCLUSION

In this Opinion, the EDPS has made recommendations to ensure compliance with the Regulation. In light of the accountability principle, the EDPS expects the ERCEA to implement the above recommendations accordingly and has decided to **close the case**.

Brussels, 27 October 2021

(e-signed)

Delphine HAROU

¹⁸ Annex 3, page 11

¹⁹ Annex 3, page 10

²⁰ page 12