Report from the Joint EASE/European Commission Discussion Forum
Brussels, 11 October 2019

GDPR, Privacy and Integrity in Journal Publishing: Good Practice for Submission and Peer Review Processes

The initiative for the Discussion Forum came from a question that was posted on the EASE mailing list: an editor was told that the journal must switch from single-blind to double-blind peer review because authors' names should not be revealed to reviewers because of the European data privacy law (General Data Protection Regulation, GDPR). The question created an intense discussion with contradictory views. Personal data protection is an important part of research ethics, so EASE consulted with the Ethics and Research Integrity Sector, DG RTD of the European Commission (EC). As many issues related to data protection during manuscript submission and review are not clear – both to journal editors and GDPR experts– EASE and the EC brought together different stakeholders in research and publishing to discuss how GDPR affects communication of research results and peer review.

The Discussion Forum started with the presentation from Dr ***Albena Kuyumdzhieva*** from the Ethics and Research Integrity Sector, EC DG RTD, who talked about the aim of GDPR and methods of compliance. Dr Kuyumdzhieva presented the principles related to processing personal data: Lawfulness, fairness and transparency, Purpose limitation, Data minimization, Accuracy, Storage limitation, Integrity and confidentiality, and Accountability. These principles need to be spelled out in a data privacy policy of scientific journals, which collect personal data about authors and reviewers for editorial purposes. These principles are important to ensure the fair use of data, which goes beyond legal regulations as it is linked to ethics.

Mr ***Emanuele Barbarossa*** from the Open Science, EC DG RTD, talked about how GDPR affects Open Science and the European Open Science Cloud. Open Science is important for research as a way of doing science in a more collaborative, open, data-intensive and networked way. Open Science increases the efficiency, verifiability, transparency and interdisciplinarity, allows re-use of information and enables society and citizens to access scientific findings. EC is developing adequate enablers for Open Science for Horizon Europe, including EOSC – the European Open Science Cloud – an open and trusted virtual environment that researchers can access seamlessly through a non-exclusive gateway in research institutions. In EOSC, different participants have different roles, from Data Providers in EOSC to research organizations and end users as Data Controllers who deposit personal data.

Prof. ***Giorgio Pedrazzi*** from the University of Brescia, Italy, specifically discussed the GDPR principle of accountability in relation to peer review. Organizations collecting and processing personal data, including scientific journals, should put in place technical and organizational measures and demonstrate what they have done to protect personal data and how effective these measures are. This means that journals need to have adequate documentation on what personal data are collected and processed, in which way, for which purpose and for how long they are stored; document processes and procedures to address data protection issue when

journal systems are built (for manuscript submission and peer review); and documentation on how data breaches will be responded to.

Data subjects in peer review include the following: publisher, reviewer, author and editor, as well as reader and subscriber (e.g. library). Journals may also publish personal data of research subjects and personal data in competing interest declarations or acknowledgements. Personal data of researchers are also available in research databases (such as bibliographic databases), clinical trial registries, scientific reports, comments or responses to published articles, and in reports of misconduct investigations related to research published in journals.

It is important to know that GDPR provides exemptions for scientific research. For example, GDPR permits the retention of personal data for longer periods when they are processed for scientific research purposes and when there is legitimate interest for processing, when the processing is necessary for that purpose and when the legitimate interest is not overridden by an individual's interests, rights or freedoms. Appropriate safeguards should be in place when processing personal data, including anonymization, encryption or pseudoanonymization of data, and restriction of access to the data. One example of legitimate interest for personal data retention in journals is preventing and dealing with research misconduct.

Journals are Data Controllers in view of GDPR, and they have to address the following in a journal's data privacy policy, which should be publicly available:

1. Define the purpose of data processing;
2. Create notices to all data subjects (authors, reviewers, readers, editors);
3. Share code of conduct and best practice;
4. Establish procedures for data retention (what really needs to be kept and for how long and where);
5. Define the procedure for how subjects may request their data;
6. Define legitimate interests for keeping personal data for an extended time period;
7. Put in place security measures and ways for keeping track of data. Journals transfer some of these actions to publishers or manuscript submission systems, which then serve as Data Processors.

The conclusion from Prof. Pedrazzi's presentation that genuine protection of personal data privacy can coexist with public accountability in scientific journals was supported by Prof. ***Flaminio Squazzoni*** from the University of Milan, Italy. Prof. Squazzoni defined scientific peer review as an inherent safeguard against the misuse of research data. He described the experience from the COST Action PEERE (New frontiers of peer review) where researchers and publishers teamed up to share full data from many journals to study the review process and editorial decision making. He argued that GDPR should not and does not endanger scientific peer review and publishing. Journals necessarily balance the need to protect personal data and increase the diversity and inclusiveness of peer reviewers, protect the autonomy and independence of reviewers, and protect the integrity of the published record in scientific misconduct allegations.

The final speaker at the Discussion Forum was Ms ***Helen Gainford*** from Elsevier, who talked about a publisher's experience in identifying non-compliance issues to GDPR principles in the peer review process. In relation to the principle of *Lawfulness, fairness and transparency*,

the publishers/journals need to identify the lawful grounds for processing, create a privacy policy to explain how personal data are collected and processed, and execute privacy impact assessment where applicable. It is important to keep in mind that journals may also deal with personal data that are included in articles/data sets and should ensure the researchers have permission to publish these.

In establishing the *Purpose limitation*, it is important that publishers identify the specific purposes for processing and communicate this clearly, and that they do not use personal data for other purposes without lawful basis. Journal editors should consult with the publisher if wishing to use personal data for a purpose other than the publishing process, and should not use personal data for their own purpose unless use and method is agreed upon.
In order to respect *Data minimization*, publishers and editors should collect only the personal data that is needed to meet the purpose, regularly review data collection forms or systems to ensure that the collected data are minimal, and that it is clearly identified what is mandatory and what optional data to be collected.

For *Accuracy* in personal data collection, it is important that personal data collected are accurate and up to date.
In relation to *Storage limitation*, personal data should be kept as long as they are required for the purpose and anonymization should be used where possible. When data are no longer needed, they should be destroyed securely.
To ensure *Integrity and confidentiality*, publishers and editors should have developed security measures and training systems to ensure that all involved in publishing are aware and trained in best practices to protect personal data. Journals should have a hotline for breaches of personal data protection and procedures to deal with them.
Finally, *Accountability* is ensured when data protection by design and default is followed, when the processes and procedures are clear and available to data subjects and when legitimate interest assessments are completed.

After the presentation, the Forum participants continued the discussion in smaller groups, addressing the four issues related to the implementation of GDPR in the peer review process and provided recommendations to scientific journals.

- *One size fits all? GDPR and different journal and publishing models*
Is it impossible to have a single peer review system for all journals, nor it is required by GDPR. Journals are entitled to run whichever peer review system suits their own community and journal. GDPR does not affect the review workflow but must be taken into account when dealing with any personal data. Editors must make the peer review system entirely clear to all parties and not assume that authors or reviewers know what “double blind” or “transparent” review means. This may require additional wording in invitation emails, etc. to ensure all parties know what personal information will be disclosed and to whom (e.g. reviewer names to authors). Personal data captured during the editorial process (author and reviewer details) should be maintained in a secure database, and not reused for any purpose other than the journal operation. There is no legal reason why these data cannot be retained: GDPR allows for data to retained and used for “legitimate interest” – in the editorial environment this means capturing and using data for editorial purposes. Each journal can capture personal data (e.g.

reviewer names and emails) and use them for review invitations, with the caveat that the individual can request removal.

- *Reconciling the right of a data subject to demand removal of their data with the need to keep the data*

There was great concern about the impact on the scholarly record if individuals are able to demand removal of key data. If data are removed from data sets, that would undermine the validity of the published research and it would be impossible to obtain clearance to reuse data for future purposes. On the other hand, there is also increasing ability to identify individuals from "anonymised" data and images through sophisticated artificial intelligence programs. If individuals demand removal of their data from the reviewing history (e.g. reviewers' names), then the historical record would be undermined. This could lead to problems in investigating misconduct if an individual's identity has been removed. However, the major submission systems allow for personal data to be removed by anonymising records, but the record itself cannot be removed.

There is also concern about patient consent forms. One opinion was that the patient form must not be sent to the editorial office as it is likely to contain personal data that the patient had not agreed to share. Editors should only ask the authors to warrant that they have obtained patient consent. The other opinion was that authors were not always clear about consent forms, and muddled them with medical consent forms. It was therefore felt important for the editorial office to receive and approve consent forms – but not necessarily retain them, they can be destroyed. However, if this system is to be used then consent forms must include permission to share the form with an editorial office.

- *Boundaries of the author's right to access to data*

An author or a group of authors submitting a manuscript to a journal provide their personal data and some administrative data. Usually one person provides personal data on behalf of co-authors. Authors can also provide personal data, such as an e-mail address, for potential reviewers. The authors should have the right to access their data in the journal's records. The boundaries to an author's right to access data include the protection of rights of all others involved in the assessment and publication process, and confidentiality of the peer review process. Exemption from these boundaries is possible only with the consent from all the parties involved.

- *Responsibility for ensuring that no personal data are contained in the published article*

Implications of online publishing and open access for personal data relate specifically to images for which patients had granted permission to be used for medical publishing. They may not be aware that the image can be reused separately from the article, for example on public websites unrelated to medical research. This is increasingly likely to happen with open access publishing under CCBY licences which permit full reuse rights, and the limitations of not infringing moral rights gives little protection.

EASE hopes that this brief overview will help journal editors understand the challenges imposed by the GDPR on the peer review process. EASE recommends that journals create or adjust their policies and practices to make sure that all involved in the peer review process are able to identify their rights and any restrictions to them.

This summary is from a Discussion Forum held in Brussels on 11 October 2019 organised jointly by EASE and the Ethics and Research Integrity Sector, DG RTD of the European Commission.