GDPR Guidebook for PKP Users

Version 1.0, Published April 30, 2018
Coordinated by James MacGregor,
Associate Director, Strategic Projects & Services
Public Knowledge Project
support@publicknowledgeproject.org

Contents

Introduction

This guide provides advice for users of PKP applications on how to approach the EU General Data Protection Regulation (GDPR), which goes into effect on May 25, 2018. It provides guidance on how to best configure OJS to be GDPR-compliant and includes information on some of the policies that those using OJS with EU clients will want to consider.

PKP's approach to the GDPR is an ongoing engagement, and will include code changes within PKP applications, and revisions to this guide, over time. Some information on near-term application changes are included at the end of this document but should not be considered exhaustive.

PKP, as a software provider, has a responsibility to provide secure software and timely bug fixes, and we welcome this opportunity to strengthen privacy rights. Those that host the software, as well as those who utilize it to publish journals, books and other artifacts, should consider ways of addressing the rights and responsibilities involved in scholarly publishing.

This guide should not be considered a source of legal advice. It is not a substitute for consulting appropriate legal authorities in your jurisdiction. It is up to publishers to determine whether and in what ways the GDPR applies to their publications. For further information on PKP's approach and to discuss aspects of this guide, please contact <support@publicknowledgeproject.org>.

PKP would like to thank our development partners for their many significant contributions to this document, in particular:

- Dulip Withanage, Heidelberg University
- Antti-Jussi Nygård, journal.fi
- Svantje Lilienthal, Free University Berlin

Key Terms

Consent: the agreement of a data subject to share personal data. In order to satisfy GDPR, consent must be unambiguous (and in the case of sensitive personal data must be explicit, i.e. “opt-in”), and must be able to be withdrawn.

Data Controller: the entity that dictates the terms for processing data. With respect to PKP applications, this would be the editorial management team.

Data Processor: the entity that manages all processing of the data on behalf of the controller - typically the journal, conference or press manager in combination with any systems administrators and service providers.

Data Subject: a natural person whose personally identifying information may be tracked within a given system.

General Data Protection Regulation (GDPR): The EU’s new comprehensive set of regulations for the handling of personal data on the Internet by service providers. It goes live on May 25, 2018 and is pertinent to anyone who manages personally identifying information of EU citizens. The complete regulation is available here: https://www.eugdpr.org/. The GDPR defines the responsibilities that Data Controllers and Data Processors must adhere to with respect to the collection, processing, storage and destruction of any Personally Identifying Data that can identify a Data Subject.

Lawful Basis for Processing Personal Data: the basis by which a data controller must explain their ability to process data. The most common lawful basis is by consent.

Personally Identifying Information (PII), or **Personal Data**: any information that can potentially be used to identify a person, such as: their name(s); email address; mailing address; phone number; social network posts; or an IP address.

Publisher: For the purposes of this policy and document, publisher refers to those responsible for the scholarly publication, be it a journal, book or other artifact, and may, in the absence of a formal publisher, refer to the editor-in-chief or the editorial team behind a single independent journal.

Rights of the Individual (Data Subject): The GDPR mandates the following rights of the individual, which it refers to as the “data subject”:

- the right to be informed;
- the right of access;

- the right to rectification;
- the right to erasure;
- the right to restrict processing;
- the right to data portability;
- the right to object;
- the right not to be subject to automated decision-making including profiling.

In order to adhere to the GDPR, people acting in the role of data controller, in conjunction with those serving as a data processor, must provide adequate means for individuals to assert these rights.

What’s the Deal?

The EU has developed and, on May 25, 2018, will be making active, a comprehensive set of regulations dictating how personal data is to be managed. These regulations dictate the rights that individuals have over their own personal data.

Anyone who interacts with EU citizens on the web must take the GDPR into account. If you are a publisher that accepts EU user registrations, EU reader subscriptions, or even EU-based web visits, this means you!

The processing of Personally Identifying Information (PII) should only be undertaken on a lawful basis, which typically involves the consent of the participant. In cases where general PII is processed, unambiguous consent is fine (e.g., a statement regarding cookies). In cases where more sensitive PII is processed, explicit consent must be given. Consent may be revoked by the data subject at any time. The data subject may also exercise their other rights at any time, and those acting as Data Controllers and Processors must have a means to address those requests.

Finally, Data Controllers and Data Processors have an obligation to ensure the proper storage and security of any processed PII and must also notify affected Data Subjects within established timeframes if a breach has been identified.

PKP advises four separate steps towards compliance:

1. Understand what personal data you process: what it is, how it’s stored, and how it can be accessed, modified and erased;
2. Develop adequate internal data storage and security procedures, including a security breach notification policy;
3. Develop and provide adequate data policies, including a contact mechanism, for your audience;
4. Configure your platforms to be secure, and to track the minimum amount of data possible.

Scholarly Publishing, Data Privacy, and the Public Interest

The realm of scholarly publishing raises a number of considerations for compliance with GDPR, particularly around the regulation's core principle of the right to be forgotten. One function of scholarly publishing is to produce a historical record of the process involved in reviewing, editing, and publishing research and scholarship, with its own norms of confidentiality and privacy. As such, this form of publishing falls within what the GDPR recognizes as a need "to reconcile the right to the protection of personal data pursuant to this Regulation with the right to freedom of expression and information, including processing for journalistic purposes and the purposes of academic, artistic or literary expression" (GDPR, Article 85). More specifically, the GDPR specifies that "the right of erasure" (Article 17) holds in situations in which "the personal data are no longer necessary in relation to the purposes for which they were collected or otherwise processed." In scholarly publishing, data concerning the authors, editors, reviewers, and others involved in the editorial and publishing process remains necessary for the purposes of the journal or press, and, as such, forms part of a record that the GDPR allows "for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes the preservation of which is in the public interest" (Recital 65). This does not apply to everyone registered with PKP software, as someone solely involved in the role of a "reader" would have reasonable grounds to expect a request to be forgotten to be honored by a journal or a press.

What Data do PKP Applications Process?

PKP applications process personal data as a fundamental part of their operations. Most data is only provided by consent, ie. through manual user registration, though some visitation data (eg. cookies, usage logs) may also be recorded.

User Registration Data

When a visitor creates a user account in a PKP application, the following personal information is processed and stored (with some minor variation between OMP and OJS, and from version to version):

- Salutation
- First name*
- Middle name
- Last name*
- Suffix
- Username
- Gender
- Password (encrypted)
- Email address*
- ORCiD ID
- Website
- Mailing Address
- Country
- Phone
- Fax
- Affiliation
- Biography
- Registration date
- Last login date
- Locales
- Reviewing interests
- Role registrations (author, reader, and/or reviewer)

Only the username, first name, last name, email and password fields are required.

Storage

This information is stored in the application database. Only the user password is encrypted.

Availability and Access

This information is available to the user via their User Profile (and, with the exception of the username and dates, can be edited). System administrators, journal managers, and editors can also access and edit this data (except the username and dates) via the application back end. The data can be downloaded by journal managers in XML format. The data is not otherwise publicly available.

Erasure

This data can be erased by the journal manager using the Merge Users tool, without affecting any editorial records. The erasure is subject to the considerations raised in the section “Scholarly Publishing, Data Privacy, and the Public Interest”, above.

Contributor Metadata Information

When a manuscript is submitted to a PKP application, contributor information is included. Contributors can be authors, translators, volume editors, and so on. This information is stored as submission metadata and is provided as part of any published manuscript record. The following contributor information is collected:

- Salutation
- First name*
- Middle name
- Last name*
- Email address*
- Suffix
- ORCiD ID
- Website
- Country*
- Affiliation
- Biography

Only the first name, last name, email address and country fields are required.

Storage

This information is stored in the application database.

Availability and Access

This information is available to almost any submission participant, with some restrictions to preserve the blind peer review process. In short: contributing

authors, editors and editorial assistants can all see this data; in most cases, only editorial staff can edit this data after submission.

Editors can download this data via author and submission reports.

Most importantly: once a submission has been published, this data is made publicly available online in a variety of ways. It is available on the submission home page to readers, is available to indexing services in underlying metadata tags, is available via an OAI-PMH endpoint for harvesting, and may be made available in any number of ways via other system plugins.

Erasure

This data can be erased by any editor by editing a submission's metadata. This can be done at any point of the submission process, including after publication. Erasure is subject to the considerations raised in the section "Scholarly Publishing, Data Privacy, and the Public Interest" above.

Workflow Data

All PKP applications track workflow information, mostly as submission-specific editorial history. The system tracks:

- All actions taken on a submission, and by whom;
- All notifications sent regarding a submission (including who sent and received the notification);
- All reviewer recommendations;
- All editorial decisions;
- All files uploaded as part of the submission process, including files that may have personally identifying information in the form of file metadata or in the files themselves.

Storage

This information is stored in the application database, with the exception of any uploaded submission files, which are stored in the application's submission files directory on the web server.

Availability and Access

Submission participants have access to different amounts of workflow data depending on their role. Journal managers and editors can access all submission data; section editors and editorial assistants can access all submission data only for those submissions to which they have been assigned; authors have limited access to their own submissions and are only able to see

the data they have supplied, or that editorial staff have explicitly made available to them.

Erasure

This data can only be erased by the editor, by rejecting and deleting the submission outright; or by a systems administrator via direct intervention into the underlying database or submission files directory. Erasure is subject to the considerations raised in the section “Scholarly Publishing, Data Privacy, and the Public Interest” above.

General Visitor Information

PKP applications also collect general visitor usage data, including:

- Cookie information, to manage session history. Cookies are required to maintain a login session in PKP applications.
- Optionally, detailed usage log data, including: IP address; pages visited; date visited; and browser information, in application log files, as part of the Usage Statistics plugin. An anonymization option is available to privatize this information.
- Optionally, country, region and city information, in the metrics database. This data collection requires additional setup and is not enabled by default.

Other data may be tracked, either on the server or via third parties:

- Script loads from CDN servers;
- IP address information (including date, browser, etc.) in web server logs (separate from application log files as part of the Usage Statistics plugin).

Detailed instructions in limiting the amount of data you collect, and providing consent for the data you collect, can be found below.

Storage

- *Cookies:* A cookie (usually titled “OJSSID” or “OMPSID”) is created when first visiting a PKP application and is stored on the visitor’s computer. It is only used to store a session ID, and to facilitate logins. (If the visitor blocks cookies, OJS will still work properly, though they will not be able to log in.)
- *Usage Statistics log files*: As part of the usage statistics framework and plugin, OJS *may* store detailed application log files in the submission

files directory (configured as the files_dir parameter in the OJS config.inc.php file), in a “usageStats” directory.
- *Geographical data*: Filtered usage data, including possibly geographic data, is also stored in the OJS database, in a “metrics” table.

Availability and Access

- *Cookies*: These are available via the visitor’s browser settings.
- *Usage Statistics log files*: Only individuals with server file access can access application log files.
- *Geographical data*: Journal Managers can access filtered usage data by using the OJS usage report plugins.

Erasure

- *Cookies*: These can be deleted via the visitor browser.
- *Usage Statistics log files*: These can be erased by system administrators with file access.
- *Geographical data*: This can only be erased by deleting records from the database directly, which also typically requires system administrator access.

What Policies Should Publishers Provide (and Where?)

The following policies should be provided in clear and plain language. Most policies can - and should - be added to the journal's About page. In the case where there is a dedicated field to do this (e.g., the journal's Privacy Statement in OJS 2, and OJS/OMP 3), do so there. If there isn't a dedicated field, these policies can be added like so:

- *OJS 2*: Journal Management > Setup > Step 2.5: Add Item to Appear in "About the Journal"
- *OJS/OMP 3*: Settings > Journal > Masthead > About the Journal.

It is also worth adding a link to these policies in email footers/signature fields. A simple statement like "Journal Data Policies: <url to About page>" would suffice, but each individual policy could also be linked separately. Email footers can be edited in the following location:

- *OJS 2*: Journal Management > Setup > Step 1.4: Email Identification, in the Signature field.
- *OJS/OMP 3*: Settings > Workflow > Emails, in the Signature field.

1. Consent policy for registration and contributor data collection

Users registering an account should be provided with access to the publisher's data management and privacy policies and should provide explicit and unambiguous consent that they understand these policies and that their data is being processed. Similarly, authors should consent to having their personal information processed by the system. They should have a means to remove their consent.

OJS 2
OJS 2 does not provide a means to add a dedicated consent statement to user registration or author submission. One possible workaround for this would be to adapt the Privacy Statement (Journal Management > Setup > Step 2.3) to a more general Privacy and Consent Policy, in which the journal's privacy approach is outlined and usage of the website (including registration and author submission) is acknowledged. This Privacy Statement is included as part of the user registration and author submission pages and is also made available under About the Journal.

A consent policy can be added to Journal Management > Setup > Step 2.5: Add Item to Appear in About the Journal. This will be made available in About the Journal.

You may also want to add a consent statement to your Author Submission Checklist, in Journal Management > Setup > Step 3.1 Author Guidelines.

OJS/OMP 3
These applicaitons do not currently provide a means to add a consent checkbox item to user registration, but this has been filed in GitHub and will be available in OJS/OMP 3.1.1-1, on or before May 25.

You may also want to add a consent statement to your Author Submission Checklist, in Settings > Workflow > Submission > Submission Preparation Checklist.

2. Privacy Policy

The Privacy Policy, which can be edited in application settings and subsequently made available on the publisher's About page and at various specific stages of the registration or data submission process, sets out the publisher's commitment to protecting the privacy of its users while adhering to best publisher practices. Note that this statement should include reference to (a) the privacy protection afforded by the software, which can only be maintained if those responsible for the software ensure that the hosting service is using the latest application version available for the software; and (b) the Privacy Policy of the publisher's online hosting service for the publication, which needs to be referenced explicitly. The privacy policy should provide a means by which data subjects can contact the data controller to exercise their Rights of the Individual as defined above.

This setting can be configured in the following locations:

- *OJS 2*: Journal Management > Setup > Step 2.3: Privacy Statement.
- *OJS/OMP 3*: Settings > Workflow > Submission > Privacy Statement.

Sample Data Privacy Policy

The data collected from registered and non-registered users of this journal falls within the scope of the standard functioning of peer-reviewed journals. It includes information that makes communication possible for the editorial process; it is used to informs readers about the authorship and editing of content; it enables collecting aggregated data on readership behaviors, as well as tracking geopolitical and social elements of scholarly communication.

This journal's editorial team uses this data to guide its work in publishing and improving this journal. Data that will assist in developing this publishing platform may be shared with its developer Public Knowledge Project in an anonymized and aggregated form, with appropriate exceptions such as article metrics. The data will not be sold by this journal or PKP nor will it be used for purposes other than those stated here. The authors published in this journal are responsible for the human subject data that figures in the research reported here.

Those involved in editing this journal seek to be compliant with industry standards for data privacy, including the European Union's General Data Protection Regulation (GDPR) provision for "data subject rights" that include (a) breach notification; (b) right of access; (c) the right to be forgotten; (d) data portability; and (e) privacy by design. The GDPR also allows for the recognition of "the public interest in the availability of the data," which has a particular saliency for those involved in maintaining, with the greatest integrity possible, the public record of scholarly publishing.

3. Cookie policy

OJS uses cookies to manage user sessions (for which they are required). Cookies aren't required to simply visit the site and read content. A cookie policy should be available from the About page in clear, precise language, and you may also want to provide a common "pop-up" cookie alert. Sample cookie policy language can be found online, for example here. A cookie policy and alert can be added to PKP applications in a few different ways:

OJS 2

1. Add a cookie policy in Journal Management > Setup > Step 2.5: Add Item to Appear in "About the Journal". This will be made available in About the Journal.
2. Install the Cookie Alert plugin at https://github.com/ictineo/ojs-cookiesAlert and configure it with an appropriate consent notice. Make sure to provide a link to the larger cookie policy

OJS 3

1. Add a cookie policy to the About section in *Settings* > *Journal* > *Masthead* > *About the Journal.*
2. Use the Custom Header plugin (available in OJS and OMP 3.1+) to add cookie alert popup code. This code can be found at various websites online. See https://cookie-script.com/ for one example. (Note: a future version of OJS will have an explicit cookie consent option. See this feature request for more information.)

OMP 3

1. Add a cookie policy to the About section in *Settings* > *Press* > *Masthead* > *About the Press.*
2. Use the Custom Header plugin (available in OJS and OMP 3.1+) to add cookie alert popup code. This code can be found at various websites online. See https://cookie-script.com/ for one example. (Note: a future

version of OMP will have an explicit cookie consent option. See this feature request for more information.)

Configuration Recommendations for GDPR Compliance

The following practices will provide a higher level of data privacy support, and are recommended as part of a reasonable attempt to fulfil GDPR requirements.

Use SSL/HTTPS for all web traffic. OJS and OMP can be used in conjunction with an SSL certificate so that all traffic between the user and the server is encrypted and transferred via HTTPS. In order to enable this, install an SSL certificate for your domain (or ask your service provider to do so) and set “force_ssl” to “on” in your config.inc.php file.

Disable CDN usage. Content Delivery Networks (CDNs) are used by OJS and OMP to deliver some content, including javascript and fonts. Any CDN can record and track detailed visitor information whenever they are loaded by the web browser, including time; user IP address; web browser; and page loaded. CDNs can be disabled in config.inc.php by setting the “enable_cdn” to “off”. (Note: this will not necessarily disable CDNs are added to the system in other ways, such as via code customizations, via third-party plugins, or in form fields.)

Restrict usage of other third party scripts. Third-party scripts, such as Google Analytics, should only be used if the application is required and the implications are understood. The use of these scripts should be properly identified in the Privacy Statement.

Anonymize usage data. OJS and OMP both have a Usage Statistic plugin that provides detailed metrics on page views and galley file downloads. It also creates and stores log files containing detailed information including IP address, date/time visited, page views, and browser information. This plugin does have a “Respect data privacy” option that will hash IP addresses, and inform visitors that this data is being tracked (with an option to opt-out). More information is available in the following locations:

- OJS 2: Journal Management > System Plugins > Generic Plugins > Usage Statistics Plugin > Settings.
- OJS/OMP 3: Settings > Website > Plugins > Generic Plugins > Usage Statistics Plugin > Settings.

Enabling the “Respect data privacy” option will require direct system administrator assistance.

Use the Sharrif Plugin for sharing/social media. Social media platforms like Twitter and Facebook all provide ways to embed sharing options and other social features into your sites, but similar to CDNs and other third party script options, these embeddable scripts typically allow the social media platform to track usage of your website. OJS-de, the German OJS user network, has developed a plugin to provide social media and sharing functions using the privacy-respecting Sharrif solution. It is available here, for OJS 2 and 3, and OMP 3: https://github.com/ojsde/shariff/releases.

FAQ

What about data collected prior to May 25, 2018?

- Policies and mechanisms should still be enforced.

Can editorial history be removed at the request of a user?

- See “Scholarly Publishing, Data Privacy, and the Public Interest” above.
- PKP will monitor GDPR activities to see whether and how this data should be treated as removable, or anonymizable, without compromising the integrity of the editorial process.

What about server log tracking (IP addresses, etc.)?

- This would be the responsibility of the service provider (who would be the data controller) in cooperation with the publisher (who would be the data processor). The service provider should also be GDPR compliant.

Are there any planned updates to OJS and OMP to further protect data privacy and support the GDPR?

- Yes. More information can be found in our GitHub Project page: https://github.com/pkp/pkp-lib/projects/11. Some items will be added to OJS/OMP 3.1.1-1, to be released before the May 25 GDPR deadline, and other, less critical items will be added over time.
- … and, no. At least, not for OJS 2.x or OMP 1.x. All updates will be made to the stable OJS and OMP 3 lines only.

You are releasing code for OJS/OMP 3.1.1-1. What if I can’t upgrade between the release date and May 25?

- Follow the general directions above to the best of your abilities, and plan on upgrading as soon as you can. Our understanding is that there will be some leeway given to organizations who can demonstrate that they are proactive in implementing GDPR solutions but who may not be 100% ready by the May 25 deadline.