1 Disclaimer

This document was written by the DSpace GDPR working group. None of its members is a lawyer. This document is written on a best effort base and must not be used in place of a legal advice. Therefore this document should be considered incomplete and questionable.

2 Introduction

The General Data Protection Regulation (GDPR) has come into effect since May 2018. GDPR gives EU citizens control over their personal data, by ensuring that companies in Europe and worldwide will process the personal data of European citizens carefully and following best practices and strict rules.

Before the GDPR became a law, the European Union had the Data Protection Directive 95/46/EC (DPD) in place. Every country within the European Union had to adopt the DPD into national laws. That lead to a situation where every country had different national laws regarding data protection. While GDPR should harmonize the laws, also today the rules in different European countries are different in their strictness. Data protection rules from Spain are not the same as in Germany. Nevertheless the GDPR is the minimum every European country has to follow. The GDPR is also important for institutions outside of Europe as GDPR regulates data protection for any service that can be used from the European Union, so in effect any service that is globally reachable over the internet.

While the GDPR is written law, laws are always open to interpretation. At the time the DSpace GDPR Working Group wrote this document, certain questions are still undecided and even lawyers pointed out that some questions won't be clarified before court sentences are made.

3 What does GDPR mean for DSpace?

An installation of any release of DSpace published until the writing of this document (February 2020) is not compliant with GDPR and any organization using it risks significant financial penalties. For DSpace to become compliant with GDPR technical and organizational issues have to be resolved.

3.1 Data we do not consider to need protection

GDPR includes exemptions in relation to specific personal data processing activities. There are two important exemptions within the framework of research activity: on the one hand, article 85 of GDPR refers to legitimate exemptions from several chapters (chapter II-VII and IX) if they are necessary to reconcile the right to the protection of personal data with the freedom of expression and information. On the other hand, article 89 of GDPR derogations apply in relation to processing for archiving purposes for the public interest and for scientific and historical research and statistical purposes. As a consequence, considering these two articles, publication of names of authors of research articles and any other result of research activity on a repository cannot be considered a GDPR infringement.

With regard to DSpace that means that all bibliographic meta data can be considered public information that do not need protection. We can expect that an author of an article wants to be named as such. While looking on the metadata schema fields DSpace includes by default, only `dc.description.provenance` seems to contain personal data that has to be protected. All other metadata fields can be considered to be public and do not need to be protected. Within the technical metadata like e.g. the submitter of an item, more data might need protection as discussed bellow.

3.2 Data that we consider personal data

DSpace stores data about people that work with the system, be it submitters, editors, admins etc., and about the work they do. This is to be considered personal data. The data is not only stored as metadata or other database records, but also in log files and in Solr records.

3.2.1 Data stored within DSpace's database

- user profiles (email address, telephone number, given, and sure name, preferred language, date of last login)
- a link between every item a person submitted and the person's DSpace account
- to every item its provenance history:
    - the name and email address of the submitter, the date and time when it was submitted, and the bitstreams it contained at that point in time. This data is necessary to be kept as the organization running the repository must be able to know who submitted a certain item and file for legal reasons.
    - every result of a workflow step (rejections including the reason for the rejection, as well as who let the item pass a certain workflow step)
    - every change of bitstreams (deletion and addition of a bitstream)
    - submission of items by import tools
- information about the rights that an account was granted within DSpace. This includes a list of groups an account is linked to, as well as technical information like which action (read, delete, edit, administrate, …) a user may perform on which object within DSpace.
- the email address of a person, that used the "Request a copy" functionality, his/her name, and the request message
- For self-registration and "I forgot my password" functionality, DSpace stores the email address used for the request and a unique token

- subscriptions of registered users to collections. The eperson id is used to link a user to a collection. Upon deletion of an eperson, his/her subscriptions will be deleted as well.

3.2.2 User generated / uploaded contents

DSpace stores user generated content, such as research results (publication and data) from all kinds of disciplines. In some disciplines it is highly expectable that the user generated content may contain personal data.

3.2.3 Personal Data stored in SOLR Indexes

DSpace uses 4 different solr indexes, all of which potentially contain personal data.

Main Search Index

Each document in this index represents a DSpace Object of some kind. DSpace puts into the index the eperson and group ids of the EPersons/Groups that are allowed to view those objects.

OAI Index

The OAI solr index, stores metadata of items to export over OAI-PMH. It does store metadata, that we do not consider to be personal data, except the email address of the person that originally submitted the item to the repository.

Statistics Index

The DSpace statistics index stores information about many of the usage events that can occur in DSpace:

- Searching for items
- Viewing item metadata
- Downloading bitstreams
- Submitting new items
- Workflow events like administrators selecting a pooled workflow item, returning an item to the pool, installing an item to the repository, giving back an item to the submitter for corrections, etc.

In a standard DSpace installation, all these event records are relevant pertaining to GDPR, because all of them carry identifying information:

- the IP address and hostname of the computer or router the user used to connect to DSpace
- the session id that DSpace gave to the user on his first connection
- if the user has an account in DSpace and was logged in, his account id

DSpace musst offer configuration options to prevent recording this information. This needs to be developed.

In default configuration, DSpace also tries to resolve the IP address to a coarse geographic location. This could be seen as risky, however, an IP address does not normally change its physical location in a provider network with every client it is given to, at least in IPv4.

Special consideration should be given to the recording of search queries, since not only does DSpace record information identifying the user, but also the exact query the user put in, therefore enabling administrators to link users to the topics they are interested in.

All in all, it is advisable to deactivate the recording of all identifying information in DSpace (IP address, session id, user's DSpace account).

Authority Index

The solr authority index, if authorithy control function is activated, stores metadata since it authority-function can be connected to several authority sources as internal database, Orcid-org, VIAF, LCNA, etc.... It does store metadata, that we have to consider personal data: Names, name´s variants, affiliation, PIDs (personal identifiers), etc.

3.3 Necessary development work

3.3.1 Information and consent of users

GDPR stipulates that a user has to be made aware of the privacy statements of a site, of all data processing with his or her personal data and has to agree to most of the data processing before it is performed. This includes the usage of cookies.

In the following sections we differentiate between Privacy Statements, the End user agreement, and cookie preferences. Privacy Statements of a site describe the data stored, how it is processed, informs any user about his/her rights, and declares who is responsible to protect these data. The end user agreement is a list of private data stored, and information about how it is processed, that a user has to acknowledge before he/she creates an account or that he/she has to agree up on login if it was changed since the user agreed to a previous version. With the cookie preferences users can declare which cookies they are willing to accept and which not.

Privacy Statement

The privacy statement contains declarations on all personal data stored for visitors, how it is processed, and which rights the user has (we try to give an overview of rights the GDPR grants every user in the Appendix of this document). Visitors to DSpace need to be made aware of the privacy statement of the installation. If they acknowledged to have read the privacy statements, the privacy statements don't have to be shown to this user again, until the privacy statements changes. Their acknowledgement can be stored in a cookie for visitors or a field in the user profile for registered users. The cookie can be permanent if the user agrees (in an **opt-in** manner) to make it permanent. The privacy statements also have to be linked in a place easy to find for every user anytime (e.g. in the footer of the site).

The GDPR Working Group will provide an example of a privacy statement. Nevertheless, the privacy statement will have to adopt special national regulations (see Introduction about how GDPR and national differ).

End user agreement

Before any user is registered into DSpace, the user need to acknowledge any private data stored and processed by the site and all included services (e.g. ORCID, altmetrics, Google Analytics, ...). This has to be done using simple language and in a way that anyone can understand. Once a user creates an account he/she must acknowledge the end user agreement. If the end user agreement changes, the users has to give permission again. This can be requested by e-mail or on the next login of a user. The GDPR Working Group recommends therefore, to store to which version of the end user agreement a user agreed to, so that they can be asked again once the agreement changed on their next login. Since this is relevant only to registered users, the information should be part of the user account in the database.

Cookies

Use of cookies in the system

A GDPR-compliant DSpace will have no mandatory cookies, besides one cookie needed for identifying logged-in users ("login cookie"), one for storing his cookie preferences ("preference cookie") and one for recording acknowlegment of the privacy statement ("acknowledgement cookie"). These cookies must not be used for any other purposes. Nevertheless, more additional cookies can be added to DSpace as long as they are requesting consent on the cookies preference panel (see below "cookie consent form").

Cookies must not be permanent unless the user agrees to permanent storage. Furthermore the Working Group recommends to make cookies permanent only if it offers additional convenience to the user. Examples are the aforementioned cookies:

- Making the login cookie permanent frees the user from having to log in on every visit.
- A permanent preference cookie will make sure the user will not have to think about his cookie preferences on subsequent visits again.
- A permanent acknowledgement cookie will do the same for the privacy statement.

DSpace core functionality should be kept free of functionality that needs more cookies as much as possible, especially third-party cookies. Therefore, components of DSpace that need external services should always be optional, from the operators' perspective and also from end users' perspective. Examples of such components:

- PlumX/Altmetrics integration
- Social Media Plugins
- Statistics/Tracking tools like Google Analytics

Also, we recommend to keep DSpace free from dependencies to external services like Content Delivery Networks. This should be discussed with the DSpace Committer Group.

Cookie consent form

A GDPR-compliant DSpace must present users it considers to be first-time visitors a cookie consent form. That form will explain the login cookie (see above) as necessary cookie to provide minimum functionality of the site, the function of any other cookie, an **opt-in** choice to enable or disable those cookies individually, and inform whether any of these cookies are stored permanently. It will offer the **opt-in** choice to store the cookie preferences of the user on a permanent basis and explain that an additional permanent cookie is needed for that. The form also explains to the user how to change his/her cookie preferences and will easily be reachable in the UI.

Having temporary preference and acknowledgement cookies might inconvenience frequent visitors because they will need to go through the privacy statement and cookie preference form again. Nevertheless, DSpace will not try to use other means to prevent that inconvenience, since that would need an identification method the visitor did not agree to. Most cookie consent forms present a huge green button "allow all cookies". While we do not want to push the user towards simply allowing all cookies, we would recommend a user friendly layout and therefore see the benefit of an "allow all cookies button".

3.3.2 Logging & Statistics

Log files

DSpace currently stores personal data in its log files. This must be made configurable so that logging can be made GDPR compliant. Simply switching of recording of all personal data in logging is not a solution, since e.g. IP addresses might be needed for security reasons. Since you need to explain in the record of processing activities how

long different kinds of personal data is stored and for what purpose, we also need tools to sanitize log files, either by deleting or pseudonymizing personal data.

Solr Statistics Log

DSpace needs options to disable inclusion of personal data in Solr Statistics records. Those might include:

- netid
- name
- user uuid
- ip address
- session id
- coarse geografical location
- ....

All storage of personal data in log files and Solr documents should be disabled by default or we do need a tool that cleans the solr index as soon as the ip was used to detect the country of the request.

As far as we know, Atmire was contracted by Humboldt University of Berlin to change how DSpace uses the solr statistics index and to provide this solution to the DSpace community.

3.3.3 User account functions

Get a copy of all personal data

GDPR stipulates the right for users to demand a copy of all personal information stored in DSpace. Therefore, a function needs to be implemented that gathers all this information and gives it to the user as a download and in a readable form.

This document will contain the following information:

- content of the profile
- current set of permissions
- a list of submitted items (urls/persistent identifier/uuid and title)
- a list of items (urls/persistent identifier and title) that mention the user in its provenance information
- other info, like subscriptions to new releases or open tickets for password renewal, if applicable.

Delete all personal information

GDPR also stipulates the right to be forgotten. Therefore, all the information mentioned in the last part has to be deleted with the user account, except meta data we consider to be public information (see above) and provenience information (see below). We recommend this to be made configurable: Give the user the option to delete his account himself or give the option to admins only.

3.3.4 Metadata

Provenience information

Provenience information is covered by GDPR and other laws (e.g. it can be seen in regard to labour law since it contains information about which reviewer approved to archive/publish a certain item and can therefore be used to track the work an employee of an institution running DSpace performed). Nevertheless the GDPR Working Group

would recommend to always preserve provenience information, arguing that an institution running a DSpace repository has to keep the information who uploaded a certain file, to be able to react to any unlawful publications (regarding to copyright or other legal rules, that a publication can breach).

To respect legal rules that forbid to track employee's work, DSpace should have configuration options to enable/disable provenience recording by type of operation (submission, workflow approval/rejection, file upload, ...) and user group (submitter, editor,...). For every category, the choices should be to switch off recording completely, to record without naming the person, and to record with full info.

On a long term basis, it should be discussed if we should use other ids than email addresses to identify users/accounts in provenience information.

3.3.5 ORCID

Within ORCID a user can regulate quite granular which information can be seen by whom. DSpace uses only publicly available data. It copies those into its Authority Index in Solr (see above). We need to develop a mechanism that automatically updates these information and even deletes information if the user changes the settings about what should be publicly available and what not within ORCID. We need a cronjob that updates the information obtained by ORCID regularly.

3.4 Improving development guidelines

Development guidelines should be adapted to GDPR requirements:

- The rules to approve code changes by Commiters should be extended to check code changes in regard of being relevant to GDPR: If additional data is gathered by a code change, it has to be ensured, that the data is considered in the function that exports all information about a user and in the function that deletes all information about an user. All features should extend the part of the DSpace manual about GDPR once that part is written. That documentation should include a GDPR compliant configuration example.
- It should be possible to host all resources DSpace needs on first-party servers without the need for CDNs, as that carries the risk of additional user tracking even with third-party tracking prevention in browsers.
- Wherever possible, DSpace functionality should be implemented without the need for external services. If external services are being used, an assessment of the GDPR compliance of those services needs to be done and documented.
- Features that cannot be implemented without integrating external services and which are not absolutely necessary for core functionality should offer an easy way to disable them and be disabled in default configurations.
- Generally speaking, it is not sufficient to have only the option to disable a feature altogether. Since the implementations of GDPR and other data protection legislation might be different in different countries, a finer granularity is desirable. For example, in some cases recording the coarse geographic location of the origin of a request might be ok, in others it might not.

3.5 Necessary communication into the Community

3.5.1 Raising awareness

At the current stage, it is not possible to run a GDPR-compliant DSpace installation created from the vanilla codebase. While this is delicate, we need to find a way to communicate that to the community not only in Europe, but world-wide (remember, all web services reachable from Europe have to be GDPR-compliant, not only those being operated in Europe or by a European legal entity). Also, being unaware of this does not make you any less

liable for GDPR infringements, and GDPR and its local implementations carry severe penalties for those infringements.

Creating a GDPR-compliant DSpace repository, or updating a repository to make it compliant, is not only a technical issue, but also an organizational issue. Repository operators need to be made aware of the fact that they will have to prepare lots of documents in order to be GDPR compliant, as described in the Documentation part below. Sample documents will be prepared for this, but those will need to be adapted for each repository and especially for different legislations, and also, repository operators will want to translate those documents into their native language.

The GDPR Working Group could do webinars about this topic, and a collaboration with the DSpace Marketing Working Group should be arranged to discuss more ways to inform DSpace users.

3.5.2 Documentation

The documentation of DSpace configuration options needs to include, in relevant cases, hints on how to create GDPR-compliant configurations, and where to also take other considerations into account. An example would be security considerations, where you wish to record IP adresses at least for a short amount of time.

Besides documenting all the technical provisions that future GDPR-compliant versions of DSpace will have, it is in our opinion also necessary to provide information on how to create GDPR-compliant legal documents pertaining to the operation of a DSpace instance. The GDPR Working Group already started to create examples and will continue with that. We want to add a section about GDPR to the DSpace manual gathering all relevant information.

Possible examples include:

- Policies
- Licenses
- End user agreements
- Privacy Statements
- Imprints
- Records of processing activities
- Texts for the cookie consent form depending on optional features used

3.6 Organizational issues repository operators should deal with

Repositories do need data protection policies explaining which data is gathered and stored, how it is processed, how it is protected, and which rights and possibilities data subjects do have within this regard. As part of this, repository operators should list the contact information of a person in charge to deal with requests regarding the GDPR and the specific repository. The GDPR Working Group will continue to work on recommendations about those organizational issues and will create a list in a GDPR section in the manual, that still has to be created (see above).

4 Recommendation to the DSpace Leadership Group

Some of the technical issues are planned to be covered in the fourth beta of DSpace 7. Other issues need to be developed for DSpace 7.0 or 7.1. The GDPR Working Group highly recommends to aspire that all identified technical issued should be solved with DSpace 7.1 at the latest. The GDPR Working Group cannot develop all of this functionality itself, missing the necessary development capacities. We hope that the Leadership Group and the DSpace 7 Development Working Group can help with that.

The DSpace community claims to back-port security relevant bug fixes to the three most up to date major versions of DSpace (currently DSpace 4, 5, and 6). While the GDPR is not a security relevant bug fix, the Leadership Group must discuss how to deal with older versions of DSpace. On the one hand it is important to raise an awareness within the community and to clearly document what we know, so that users can take this into consideration. On the other hand, it might bring institutions running older versions of DSpace into a problematic position, if we claim that DSpace 7.1 is now GDPR compliant up to our knowledge, which might imply that they run versions of DSpace which are not compliant. Putting older versions of DSpace on the spot of not being compliant is risky for the community. Back porting all necessary changes to DSpace 5 and 6 (4 will be eol, once DSpace 7.0 is released) would add an enormous amount of work.

The DSpace GDPR Working Group wants to work on the documentation and on examples of policies and statements as discussed above in this document. Much of this work cannot be done until the technical issues are solved, as it is not the best idea to write documentation for features that do not exist yet. Therefore the Working Group would suspend it work until some of the technical issues are solved and will resume its modus of operation when it can continue with working on the documentation. The Working Group then also wants to offer Webinars for the community, to raise an awareness for this topic and to help repository operators to achieve GDPR compliance asap. The Working Group would be grateful for any ideas how it could help the community in regard to GDPR compliance.

While the GDPR Working Group tried to understand GDPR and its effects on DSpace as good as possible, none if its member were a lawyer. It should be considered to approve a budget to the Working Group to discuss the results of the Working Group and the recommendations to the community (like an example for a data protection policy for a DSpace installation) with a lawyer.

5 Addendum

5.1 Terms & Definitions

Binding corporate rules:

Personal data protection policies which are adhered to by a controller or processor established on the territory of a Member State for transfers or a set of transfers of personal data to a controller or processor in one or more third countries within a group of undertakings, or group of enterprises engaged in a joint economic activity.

Consent

Any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her.

Data controllers

The natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law;

Data processors

A natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller; In many cases, the data controller and the data processor will be the same entity. In the example above, the organisation that provides the online form will be a data processor because the act of collecting data is included within the definition of 'processing'. A single data controller may have several data processors.

Data subject

"An identified or identifiable natural person". There is no restriction on their nationality or place of residence, however, so a data subject can be from anywhere in the world – the Regulation does not distinguish. Equally, however, a data subject has to be a person; a corporation or other entity cannot be a data subject, and information on those subjects has no protection under the Regulation.

Filing system

Any structured set of personal data which are accessible according to specific criteria, whether centralised, decentralised or dispersed on a functional or geographical basis; This is used as a generic term to cover all methods by which personal data can be collected, stored, transmitted and processed.

Personal data

"Personal data" means any information relating to an identified or identifiable natural person ("data subject"); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;

Of specific note here is that the set of characteristics above is not exhaustive: any information that could be used to identify the data subject is personal data, and this information can be in any format. This can encompass photographs, correspondence, physical media and so on.

Processing

"Processing" means any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage,

adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction;

Profiling

Any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements;

Data subjects must always be informed if any profiling processes will be performed on their personal data before they consent.

Pseudonymisation

The processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person;

While the Regulation generally considers pseudonymisation to be a positive thing, it does also specify that pseudonymised data that can be "attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person". As such, any organisation that uses pseudonymisation to protect personal data should ensure that it is not possible to identify the original data subject if additional information is made available. As noted in the definition, this should include measures to completely separate pseudonymised data from all other personal data.

5.2 Principles of GDPR

Article 5 of the GDPR outlines the six principles that should be applied to any collection or processing of personal data.

1. Personal data must be processed lawfully, fairly and transparently.

2. Personal data can only be collected for specified, explicit and legitimate purposes.

3. Personal data must be adequate, relevant and limited to what is necessary for processing.

4. Personal data must be accurate and kept up to date.

5. Personal data must be kept in a form such that the data subject can be identified only as long as is necessary for processing.

6. Personal data must be processed in a manner that ensures its security.

5.2.1 Applicability

The GDPR applies to organizations within the EU, to any external organizations that are trading within the EU, and to any service that can be accessed from within the EU.

5.2.2 Data subjects' rights

The GDPR considerably increases the rights of data subjects, such as Consent, Right to be forgotten and Data Portability. Much has been made of this – especially the 'right to be forgotten' – but the Regulation does attempt to balance those rights against the right to the free flow of information in order to support "the pursuit of economic

activities". The expanded rights granted to data subjects can generally be characterised as giving them more control over their data and giving them a better understanding of what is being done with it.

5.2.3 Consent

Data subject's consent is required to process their data. There are specific circumstances in which consent is not strictly necessary, these generally revolve around legal requirements (such as in compliance with another law, or in order to protect the rights of a data subject), or where the data subject's consent is provided through a contract they have with a third party. Beyond these sorts of exemptions, we will need to ensure that we secure consent for processing any data subject's personal information.

The specific circumstances in which consent is not strictly necessary are often mistaken. Most people thinking to have a justifiable reason to process personal data without explicit consent, have wrong expectations about what justifiable reasons are. More and more fines are imposed in significant financial amounts for intended but also for unintended misbehavior.

Data controllers will have to ensure that they secure clear and unambiguous consent from the data subject before processing personal data. Critically, the controller is not permitted to count "Silence, pre-ticked boxes or inactivity" as consent. Furthermore, processing cannot proceed unless the data subject has consented to every processing activity – if we wish to carry out six different actions with the subject's data, for instance, we need to ensure that the subject has consented to all of them.

The Regulation notes the following:

1. consent can be provided electronically using a tick-box (although, as noted above, the data subject will have to manually tick the box themselves), which is in line with the way many organisations already ensure appropriate consent for specific activities.
2. The consent document should be laid out in simple terms. In the words of the Regulation, "the request must be clear, concise and not unnecessarily disruptive to the use of the service for which it is provided". This final point may be problematic, especially where consent is required for a variety of activities.
3. Documentation of consent is crucial, and this is one key area in which legal input from your professional advisers is essential.
4. Consent can be withdrawn. Web application developers will no doubt need to design and implement robust solutions to allow data subjects to withdraw their consent in accordance with this requirement.

5.2.4 Right to be forgotten

Data subjects have the right to have any data held about them erased, if they withdraw consent of the processing.

5.2.5 Portability

Under Article 20 of the Regulation, data subjects can request a copy of any personal data held on them, and can also request that this information is transmitted to another data controller. The Regulation doesn't stipulate precisely how this information has to be presented or the format it has to be in, but it does require that it is in a "structured, commonly used and machine-readable format".

5.2.6 Lawful processing

As noted earlier, controllers are accountable for ensuring that personal data is lawfully, fairly and transparently processed. The lawfulness of processing ensures that the data subject must have given consent (thus including all of the requirements of consent noted earlier), or that the processing is necessary for certain tasks, the majority of which require consideration of the data subject's interests.

5.2.7 Retention of data

As noted earlier, data subjects have the right to be forgotten, at which point the data controller must erase all information held on them. In addition to this, however, personal data can also only be retained for limited periods, which should be clear to the data subject at the point at which they consent. This isn't a hard and fast rule, of course, as some personal data could be held effectively indefinitely (by public bodies for specific governmental purposes, for instance) and other processing, by its nature, may be ongoing. Regardless of how long personal data, are retained confidentiality and integrity must be secured – including against accidental loss, destruction or damage.

5.2.8 Data protection impact assessments

What the GDPR calls data protection impact assessments (DPIAs) are now mandatory for technologies and processes that are likely to result in a high risk to the rights of data subjects. Much like other impact assessments, we'll need to ensure that we take advice from an appropriate authority.

The supervisory authority in each EU Member State may list the specific situations for which a DPIA is or is not required. Regardless of whether this is the case, most organisations should ensure that a DPIA is part of their risk assessment process regarding personal data, and is in line with their data protection by design and by default strategies.

5.2.9 Controller/processor contracts

Where a controller contracts with a processor to process personal data, that processor must be able to provide "sufficient guarantees to implement appropriate technical and organisational measures" that processing will comply with the GDPR and ensure data subjects' rights are protected.

Many organisations will be required to appoint a data protection officer (DPO). Whether or not you need one is based on three conditions:

1. If the data is processed by a public authority or body, except for courts acting in their judicial capacity.
2. If the controller's or processor's core activities consist of processing operations that require regular and systematic monitoring of data subjects on a large scale.
3. If the controller's or processor's activities consist of processing large quantities of special categories of data and personal data relating to criminal convictions and offences.

5.2.10 Data breaches

In addition to being damaging for business, even if the authorities don't get involved, data breaches are much more strictly regulated under the GDPR. The Regulation, however, mandates informing both the supervisory authority and the data subjects themselves. Data breach reports must be made within 72 hours of the data controller becoming aware of the breach. If that requirement is not met, the eventual report must be accompanied by an explanation for the delay. The notification must follow a specific format, which includes a requirement to describe the measures being taken to address the breach and mitigate its possible side effects. Where the breach may result in a high risk to the rights and freedoms of data subjects, they must be contacted "without undue delay". This contact will not be necessary if appropriate protective measures – essentially encryption – are in place to eliminate danger to data subjects.

5.2.11 Changes to the 'Cookies Law'

The 'Cookies Law' – properly called the Directive on Privacy and Electronic Communications or the E-Privacy Directive – was controversial when it came into force in 2011 and it has remained so. The GDPR itself clarifies that a cookie could be interpreted as an online identifier, which means that it falls under personal data and, therefore, the data subject must consent. This clearly asserts that all of the cookie notifications will need to follow the normal rules for consent, and forcing the supervisory authorities to act when non-compliance is discovered.

5.2.12 IP addresses

In the same breath that the regulation declares cookies to be personal data, it also declares IP addresses to be the same. The regulation attempts to protect this information where it could be used to identify someone, but at the same time it appears to assert that an IP address could be interpreted to indicate a specific geographical location.