# EN - Policy on management, storage and deletion of customer proprietary data under The General Data Protection Regulation (GDPR)

This document describe the policies and procedures that 4Science uses for the storage and deletion of customer proprietary data, during the service period and after termination of the contractual relationship.

### Where data are stored

The data contained within the applications that 4Science provides are always saved in the cloud, on Amazon Web Services (AWS, https://aws.amazon.com ), unless explicitly specified in the contract. In the latter case, deletion procedures are also defined separately.
Such data may be contained in:

- AWS RDS Database tables (e.g.: Postgresql and Mysql) or similar (e.g.: DynamoDB, Redshift, etc.);
- AWS S3 or Glacier buckets;
- AWS EC2 EBS volumes;

Specifically for the following Saas services:

- **Saas Dataverse Services and Saas DSpace Services (GLAM e CRIS):**

1. Uploaded digital materials, accessible only from Dataverse and DSpace via specific policies set on them and from the application server, are saved on Bucket S3.
2. All files are versioned and the resulting versions are maintained for 30 days.
3. Application files are always stored on dedicated EBS volumes for the exclusive use of the customer, accessible only from the application server.
4. There are daily backups maintained for 14 days and monitored by an automatic procedure.
5. There is a database reserved for each customer (on a shared cluster) on Amazon RDS whose data can only be accessed through the application and by authorised 4Science personnel.
6. Daily backups are also made for the databases and stored for 15 days.
7. Application logs are constantly processed so as to anonymise them. They are also moved to a centralised area only for storage (30 days) and analysis of service use purposes.

- **SaaS OJS Services**:

1. Application files and uploaded digital materials are stored on dedicated EBS volumes for the exclusive use of the customer, accessible only from the application server.
2. There are daily backups maintained for 15 days and monitored by an automatic procedure.
3. There is a database reserved for each customer (on a shared cluster) on Amazon RDS whose data can only be accessed by the application and by authorised 4Science personnel.
4. Daily backups are also made for the databases and stored for 15 days.
5. Application logs are constantly processed so as to anonymise them. They are also moved to a centralised area for storage (30 days) and analysis of service use purposes.

### Security measures in case of unauthorised access

To prevent unauthorised access, AWS IAM roles and personal RSA keys are used to access the servers. Only the technical support role can access the data to verify any support requests.

1. Each machine has an access log that can be used for initial investigation.
2. AWS GuardDuty is active on the entire AWS account. It is a continuous security monitoring service that analyses and processes the following data sources: VPC logs, AWS CloudTrail event logs, and DNS logs. It uses threat intelligence feeds, such as lists of malicious IPs and domains, and machine learning to identify unexpected and potentially unauthorised/harmful activity within the AWS environment.
3. 4Science schedules periodic Application Penetration tests and Network Penetration tests (both from inside and outside the network) to ensure the infrastructure is secure.

### Monitoring tools for data security levels

Data security is ensured through versioning on S3 and backup on EBS (see "How data are stored and possibly archived"). In parallel, there is an annual procedure for checking backups. For DSpace in particular, the data are monitored by an automatic procedure to check the consistency between the S3 buckets and the hash previously saved on the database.

The technical staff of 4Science has complete access to the data in order to provide the services agreed upon with the customer. It is also possible, through the various applications, to define the level of external access to the data using specific functions.

All the services subject to this qualification collect information of a public nature, the individual institutions may decide what information to collect by explicitly setting out the terms of use and requests for authorisation.

As regards the processing of any non-public data, these are established and configured during the application setup phase.

Specifically for the following Saas services:

- **Dataverse Saas**: It is possible to define which users/groups have access to the metadata and files of a specific dataset or collection (datverse) of datasets.
- **DSpace SaaS** (GLAM and/or CRIS): The institution can determine which metadata is visible only to administrators and whether metadata of specific items should be visible only to specific user groups or after a certain period (embargo). The same access restriction can be applied individually to individual files attached to items: public, embargo, visible only to specific users/groups.
  Data are editable by individual users who create them only in a draft state, once published they can only be edited by administrator groups defined by the institution.
- **OJS SaaS**: Approved proposals/publications are by their definition of a public nature. Any access restriction can be defined by the customer for commercial reasons (sale of subscriptions/access).

Submissions are visible not only by the authors themselves but also by the editorial board according to the workflow defined by the individual journal and according to the data access policies, always determined by them and made explicit in the appropriate pages of the journal (e.g.: anonymous review of submissions, public/closed or open list of reviewers, etc.).

**Security measures in case of unauthorised access**

- If there is an internal data breach (if the breach occurs on data that 4Science processes directly): In In this case the authorized party must give timely written communication to the owner of the data, also describing what has been detected, the type of breach and the quantity of data involved, the circumstances in which the breach occurred and how 4Science became aware of it and the actions taken to remedy the situation. The notice is, therefore, timely; occurring as soon as 4Science discovers the breach. The data owner then, communicates to the Guarantor, if there are the conditions, within 72 hours from moment in which it has come to the knowledge of the violation.
- If there is an external data breach (if the breach occurs on data owned by 4Science, but entrusted to an external manager): The external data controller is obliged to notify the breach promptly and without undue delay (no later than 12 hours), either by email or by pec.

**Insurance coverage in relation to privacy risks**

The Company has a Professional Liability Insurance "Rischi Diversi" that includes, in particular, the following clauses:

- The insurance company undertakes to indemnify 4Science S.r.l. for what it is obliged to pay as civilly obliged pursuant to the law, as compensation for damage caused to third parties as a consequence of the professional activity carried out. The latter relates to the provision of integrated solutions in the field of IT, in the offer of solutions, products and/or services for security, integration and centralisation of applications; as well as in the offer of consultancy products in the area of software development, in the provision of services and consultancy for e-commerce, datawarehouse and the creation of internet applications.
- The insurance does not specifically cover damage arising from cyber crime, but a specific policy is being assessed and signed for this purpose.