

Physical Science Data-science Service: Data Protection Policy

Brian Matthews, STFC

15 April 2019

Introduction

1. The *Physical Science Data-science Service* (PSDS) is being established as a hub for collecting, accessing and reusing data for research in the physical sciences. The PSDS is formed by a consortium between the University of Southampton, and the UK Research and Innovation-Science and Technology Facilities Council (STFC) (“the partners”). The PSDS is funded by UKRI-EPSC.
2. As part of the service, the PSDS will collect and manage personal data concerning its user community. This document seeks to establish the framework for collection, storage, sharing and use of personal data for the purposes of the sound management of the project and compliance with appropriate legislation under the General Data Protection Regulations (GDPR), including the roles and responsibilities for data protection within the parties.

Personal Data Collected

3. The PSDS is intended to provide access to data of relevance to the research community within Physical Sciences. In particular, it needs to use personal data for the following purposes:
 - The PSDS acts a data broker for common access for the UK’s academic community to licenced third-party databases for the purposes of research and teaching. In this case, it needs to ensure that users accessing third-party data are from a recognized UK academic institution to ensure compliance with data licences.
 - The PSDS needs to provide metrics on usage of data, including which data collections are of particular value, and where data in particular topics may be missing. Metrics will include such information as usage from institution and usage by subject domain. These metrics will be used to report on the current status of the service to the funders (EPSC) to assess the value of the service, and to guide the future development of the PSDS to better serve the needs of the user community.

- The PSDS needs to maintain contact with the user community, both to inform the community on the status and development of the service, information on events and training, and to consult with users on the future priorities for the development of the service.
 - The PSDS seeks to provide additional services to the User Community, including: provision of training course and training material; access to personalised functionality within PSDS system, including stored searches, scratch areas for data storage, use of software tools.
4. To support these purposes, the PSDS will hold and use personal data as follows:
- A minimal collection of data sufficient to establish that the user is a legitimate UK academic user within scope of the data licence. This will use one or both of :
 - IP address to check it is within the IP address range of a UK academic institution (typically within the .ac.uk domain, although exceptions are considered on an individual basis).
 - A user registration providing a current Email address from a UK academic institution (typically within the .ac.uk domain, although exceptions are considered on an individual basis).
 - Registration of a current email address for communication and information dissemination purposes. The PSDS will verify the validity of the email address (by for example sending an email requiring a response).
 - For richer user consultation and metrics collection, and for personalised services, a *User Profile* will be maintained. Data collected in user profiles will be optional, protected with a user password, and under the user's control. This will collect further information on the user, their institution, and their research interests. This will be used to tailor the information and services offered, and to provide richer metrics on the performance and future direction of the service. We give a description of the user data that the PSDS proposes to collect in Annexe 1.
5. We do not anticipate that personal-sensitive data will be collected for any of these purposes.

Roles and responsibilities

6. University of Southampton and UKRI-STFC act as Joint Data Controllers within the PSDS.
7. Personal data will be held within the PSDS systems and services, maintained by the partners of the PSDS. Therefore, University of Southampton and STFC act as Joint Data Processors within the PSDS.
8. Responsibility for the good governance of personal data within the PSDS lies with the Investigators of the PSDS.
 - University of Southampton: Prof. Simon Coles (PI), Prof Jeremy Frey (CoI)

- STFC: Dr. Brian Matthews (Col), Dr. Juan Bicarregui (Col)
9. The Service Manager of the PSDS will have operational responsibility for maintaining the confidentiality of personal data within the normal service operations.
 10. Data protection policy and enforcement will be reviewed by the Executive Committee of the PSDS, which includes the Investigators and the Service Manager.
 11. The Investigators and Service Manager will authorize staff within STFC and University of Southampton as required for necessary service provision and development, including for user training and dissemination.
 12. The PSDS will consult with and be advised by the IT Security Officers at STFC and Univ of Southampton, and the UKRI and University of Southampton Data Protection Officers.

Data Retention, Sharing and Disposal

13. Personal data will be held within the PSDS systems and services, maintained by the partners of the PSDS, and will be shared between the partners for the sole purpose of management, promotion and development of the PSDS.
14. Primary operation of services will be within STFC's Scientific Computing Department's service operations. This is an established data centre with policies and procedures for security, monitoring and incident management. Personal data will be maintained in a database with access restricted to operations staff authorized by the PSDS.
15. The University of Southampton may establish systems to support additional services (e.g. user surveys, training data). The PSDS partners may share user data for the effective operation of these services; however, any data transfer will be kept to a minimum, with the primary source maintained at STFC.
16. PSDS may use third-parties as Data Processor for particular services (e.g. within commercial cloud service providers). However, PSDS will retain control over the user information used and not share it with the service provider. Third-parties will be within the European Union.
17. The PSDS may be required to share information on users' with data licensors. This is to establish that the users' access to data falls within the scope of the data licence conditions. This will be kept to a minimum necessary to establish a valid access (e.g. user's IP address or email address).
18. Aggregated and anonymized metrics and performance data will be shared with the project governance structures (Executive Committee, Management Advisory Panel) and with the PSDS's funders (EPSRC). Such metrics will be included in publicly available reports.

19. User information will not be shared with third party suppliers for marketing, advertising or provision of services out of the scope of the PSDS.
20. Users will have the right to edit or remove their user information and profile at any time, although records of user activity will not be removed from aggregated and anonymized metrics and performance data.
21. Users will have the option of sharing their user profile data within the PSDS for community formation and communication purposes.
22. The PSDS will perform periodic data cleansing activities, where it will establish the currency of user registrations and profiles. This will be undertaken at minimum at yearly intervals. Inactive user profiles will be registered as inactive, and deleted after a further 3 month period. Records of user activity will not be removed from aggregated and anonymized metrics and performance data.
23. The personal data will be maintained for the duration of PSDS, Jan 2019-Dec 2023, unless the service is extended with the same partners. All personal data will then be deleted after a reasonable reporting and service period.

Website and Cookies

24. The PSDS will host an open website, although some services within the website may require user registration before they can be accessed.
25. The PSDS will use Cookies deposited within the user's client for session management and collection of metrics. The use of cookies will be notified to the user on first use.
26. The PSDS may use third party cookies for the purpose website metrics and analytics (e.g. Google Analytics) to evaluate and improve the service.
27. The PSDS will not host cookies from third party supplier for marketing or advertising purposes.

Annexe 1: User Profile Definition

The PSDS will provide a facility to collect a user profile. This allows a registered user to provide more than minimal information (valid email address). The information provided will be optional for the user, protected via a user authentication system (e.g. password), and can be changed or deleted by the user at any time.

The user will have the option of sharing the profile with other users in the system to allow communication and the formation of communities of practice.

The profile will be used within the service to provide richer metrics, targeted communications, and personalized services.

Changes to the information within the profile will be communicate to the users, who will have the option of updating information as they wish.

User profile information would include:

Field	Description
Name	Name which the user wishes to be known by
Title	Title which the user wishes to be known by
Email address	Primary email address which the user wishes to register within the service. This will not be shared with other users.
ORCID ID	Registered ORCID ID, to allow linking to user's ORCID profile.
Position	e.g. Staff or student.
Institution	Academic Institution, e.g. University of Southampton
Department	Organizational unit within institution, e.g. Department of Chemistry
Areas of interest	Areas and domains of interest within the Physical Sciences. Should include terms from controlled vocabulary.