# Data Management Plan / Information Security

**Current version:**
VERSION 1.8: 20/11/2020

## Summary of terms/acronyms:

**UoN:** University of Nottingham
**EPSRC:** Engineering and Physical Sciences Research Council. The UK's main agency for funding research in engineering and the physical sciences.
**ESRC:** Economic and Social Research Council. The UK's main agency for funding research on economic and social issues.
**GDPR:** General Data Protection Regulation. A regulation in EU law on data protection and privacy for all individuals within the European Union.

## Contents:

# A. Data Collection and data access

**Data provided by a 3rd party**

1. What data is being collected:
    a. Data collection responsibilities are undertaken by the company collecting the data.
    b. For general and publicly available statistics pertaining to geographical regions (e.g. Population sizes, Indices of Multiple Deprivation, etc.), due diligence must have been undertaken by the provisioning organization.
2. Short term storage:
    a. When data is held locally, full encryption at rest will be used for all machines storing or processing data. All access will be restricted and logged. Data is held under conditions of the specific company agreements. See section D. Information Security.
3. Long term storage, metadata:
    a. When data is used to support research findings, metadata will be stored on a researcher/project/group web page[1] along with details of relevant/corresponding publications in order to support research replicability meeting UK research standards (see section C.3 RCUK, EPSRC, ESRC and Funder compliance). For confidential and/or sensitive data metadata will be limited to a researcher contact and a statement indicating that data access was under legal agreement and that "compelling legal or ethical reasons exist to protect access to the data".
4. Long term storage, data:
    a. Long term storage of 3rd party data is outside the remit of the research organization (see section C.3 RCUK, EPSRC, ESRC and Funder compliance for further details), with storage held under the conditions of the specific company agreements.
5. Data access:
    a. Data access will be limited to sanctioned members of the research team or those specified by the company providing the data on pre-authorised and correctly configured machines (see section D. Information Security).

**Personal data collected by us**

1. What data is being collected:
    a. Currently no personal data is directly collected by us. Additional documentation will be added if this changes, and data will be maintained within the provisions specified below.
2. Short term storage:
    a. Full encryption at rest will be used for all machines storing or processing data. All access will be restricted and logged. See section D. Information Security.
3. Long term storage, metadata:
    a. When data is used to support research findings, metadata will be stored on a researcher/project/group web page (UoN data store/catalog on its completion) along with details of relevant/corresponding publications in order to support research replicability and meet funder requirements. Meta-data will additionally be offered to any funder specific repositories as required, see section C.3 RCUK, EPSRC, ESRC and Funder compliance.
4. Long term storage, data:
    a. Securely maintained by the lead researcher within an encrypted container on the UoN provided filestore. In order to support research replicability and integrity, and maintain compliance with funder requirements, access details will be included in the published metadata indicating the "terms … [under which the] supporting research data may be accessed."
5. Data access:
    a. Data access will be limited to sanctioned members of the research team on pre-authorised and correctly configured machines (see section D. Information Security).

[1] In order to ensure long term persistence metadata will also be added to the University of Nottingham data store/catalog( on its completion, if such metadata is not already publically available.
https://www.nottingham.ac.uk/research/research-data-management/data-sharing-and-archiving/depositing-and-archiving.aspx

b. Data access may be granted to other research teams under funder requirements and under specific agreements when all legal and ethical issues can be mitigated.

**Non-personal data collected by us**

1. What data is being collected:
    a. Currently no non-personal data pertaining to individuals is collected by us. Additional documentation will be added if this changes, and data will be maintained within the provisions specified below.
2. Short term storage:
    a. Full encryption at rest will be used for all machines storing or processing data. All access will be restricted and logged. See section D. Information Security.
3. Long term storage, metadata:
    a. Will be stored on a UoN hosted researcher/project/group web page along with a copy of the publication. The metadata will additionally be stored along with the data.
        i. If any funder specific requirements exist (see section C.3 RCUK, EPSRC, ESRC and Funder compliance) these will be met.
        ii. Metadata will also be added to the University of Nottingham data repository.
4. Long term storage, data:
    a. If the data may be publicly released (most cases) then the data will be published online and linked via a UoN hosted researcher/project/group web page along with any corresponding publication pre-print and meta data.
        i. If any funder specific requirements exist (see section C.3 RCUK, EPSRC, ESRC and Funder compliance) these will be met.
        ii. Data will be added to the University of Nottingham data repository in order to ensure persistence.
    b. In cases where public release is not possible for legal or ethical reasons the data will be securely maintained under the same provisions as *Personal data collected by us.*
5. Data access (pre any public release):
    a. Data access will be limited by sanctioned members of the research team.
6. Data access (post any public release):
    a. Data will be publically available, but access will be logged for monitoring research impact and to maintain compliance with funding body retention requirements (EPSRC: 10 years after last access[2]), or for any other period of time stipulated by the funder.

## B. Documentation and Metadata

**Data provided by a 3rd party**

1. Documentation and metadata for 3rd party data is generally the responsibility of the providing company. Exceptions include:
    a. Metadata meeting the funder guidelines will be developed and published when data is used to support research findings. If not specified metadata meeting the EPSRC guidelines will be produced[3]. Where access to the data is restricted (e.g. due to commercial confidentiality) metadata will provide the reason for this, and summarize the conditions that might be satisfied for access to be sought.
    b. Additional documentation will be generated regarding the data when insufficient documentation is provided by the company. In this case the documentation will be returned to the company or destroyed at the same time as the data under the individual agreements with the companies.

**Data collected by us (including personal data)**

[2] Expectation VII: https://www.epsrc.ac.uk/files/aboutus/standards/clarificationsofexpectationsresearchdatamanagement/
[3] "sufficient to allow others to understand what research data exists, why, when and how it was generated, and how to access it" https://www.epsrc.ac.uk/files/aboutus/standards/clarificationsofexpectationsresearchdatamanagement

1. Metadata meeting the EPSRC guidelines by default will be developed and published when data is used to support research findings as per EPSRC expectation V & VI (See section C.3 RCUK, EPSRC, ESRC and Funder compliance). Additional metadata will be generated as required for any other funding body.
2. Documentation will be completed as appropriate. Typically this will involve data dictionaries (when data is stored in the database) and data collection descriptions. Both data dictionaries and data collection descriptions will be stored with the raw data itself or a comma separated backup of the database tables if it was directly entered into the database. This will be stored in an encrypted container (see section D. Information Security) unless the data is publicly released in which case the documentation and metadata will be provided along with the data.

## C. Ethics and Legal Compliance

**Data provided by a 3rd party**

1. Ethical and legal responsibilities are undertaken by the company collecting the data. Company specific legal and ethical procedures will be undertaken and conformed to on a case-by-case basis. Additional ethical and legal considerations, that may extend past 3rd party responsibilities, are undertaken on a per project basis within the University.

**Personal data collected by us**

1. Any personal data collection will require and be governed by University of Nottingham Ethics[4] in addition to any other funder requirements, i.e. ESRC's Framework for Research Ethics[5]. Currently we do not intend to collect any personal data. This section will be updated if this changes.

**Non-personal data collected by us**

1. Non-personal data will be collected in accordance with any applicable laws and if deemed necessary, University ethics will be acquired. When working with companies, legal and ethics approval within these companies will be acquired as required.
2. In all cases all steps to ensure EPSRC compliance will be followed (see section C.3 EPSRC and Funder compliance), in addition to any other funder requirements.

### C.1 UK Data Protection Act compliance

**Data provided by a 3rd party**

1. Data collection responsibilities are undertaken by the company collecting the data
2. As per the agreements with the individual companies, data will be altered within the system following any request, including any to ensure data accuracy.
3. Right of access to any personal data is via the *data controller*[6]*.* This is the company owning the data set. Requests will therefore be directed to the appropriate company. Any subsequent request from any company for which we hold data will be undertaken as per the agreements with these companies.
4. Data retention: Data is retained as per the original agreements by the data providers.
5. Information security: See section A. Information security and B. N/LAB Data Server
6. Data transfer. Data will be transferred as per the individual company's agreement, fully complying with the Data Protection Act

**Personal data collected by us**

1. If personal data is collected by us additional details regarding its collection and storage with regard to Data Protection Act compliance will be detailed here.

---

[4] https://www.nottingham.ac.uk/fabs/rgs/governance-and-ethics/research-governance-and-ethics.aspx
[5] https://esrc.ukri.org/funding/guidance-for-applicants/research-ethics/
[6] see http://www.legislation.gov.uk/ukpga/1998/29/section/1
and http://www.legislation.gov.uk/ukpga/1998/29/section/7

## C.2 GDPR

**Data provided by a 3rd party (that includes personal data)**

1. Data provided to us has been provided for a specific purpose (typically research) by the 3rd party who is the *data controller*[7].
2. Acting as a *data processor* we:
    a. Hold a list of all 3rd party provided data that is provided under agreement / NDA and/or contains personal information. This includes the source of the information, who we are able to share it with, the reason for having the data and how long it will be kept. In almost all cases these are part of the agreement / NDA. This list is held and maintained by the Data Protection Officer (see Section E. Responsibilities).
    b. All access to data is both restricted and logged as is the transfer of data between machines. The restrictions and specifics regarding logging are detailed in section D Data Security Summary. Data is held under conditions of the specific company agreements or at the standard detailed in the document, whichever is higher.
    c. N/LAB is governed by the University of Nottingham privacy policy, which is available online, under which N/LAB's activities are governed.
    d. A Data Protection Officer for N/LAB has been appointed who is part of the decision making team and responsible for inducting new team members. This is in addition to the University's Data Protection Officer. See Section E. Responsibilities.
    e. Have a detailed Data Security Plan (see separate document) which is updated as required to ensure technical security is up to date.
    f. Commit to holding a list of sub-processors and updating our privacy policy if we ever require one.
    g. Will report any data breaches involving personal data to the data controller and the local authority within 72 hours. Since all data is encrypted, reporting will not be required to the person other than in exceptional cases or at the request of the data controller.
    h. Personal data kept is reviewed regularly and deleted when it is no longer required or, if earlier, as specified in the individual data agreement with the 3rd party.
    i. Based on the use of whitelisted machines and a central database/server in a controlled environment (see Section D: Data Security Summary) processing, including the ability to stop, update and answer data requests can be easily done based on requests from the data controller
    j. Transfer data only inline with our data access agreements with the 3rd party. In addition we commit to only transfer data under these agreements to entities which offer an appropriate level of protection.

**Personal data collected by us**

1. Currently no personal data is collected by us. If this did change we would store and process the data under the same process as data provided by a 3rd party. In addition, acting in this case as the *data controller* we will:
    a. Update our privacy policy to explain why we need to process the personal information, re-confirming that it is clear and understandable. On any updates, individual's part of any collected data will be informed.
    b. Ensure that if any (none is currently intended) external data processors were employed there was a contract in place.
    c. Ensure consent is obtained before processing the collected data
    d. Ensure individuals could easily contact us to comment, object, withdraw consent or otherwise inform us with regard to their data and its processing.
    e. Ensure children's personal data is collected and processed only after age verification and consent from their legal guardian
    f. Continue to regularly review these documents (see section E: Responsibilities)
    g. Continue to acknowledge the need for DPIA if high-risk processing of sensitive data, something that is not currently undertaken nor planned.

[7] https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/key-definitions/

## C.3 RCUK, EPSRC, ESRC and Funder compliance

**Data provided by a 3rd party**

1. Data retention: Data is retained as per the original agreements by the data providers. This is compliant with EPSRC guidelines[8], specifically: *"Research Organisations are not expected to assume responsibility for the preservation and management of third party research data not generated within their own organisation" [clarification of Expectation VII]*. This is consistent with the RCUK Common Principles on Data Policy[9] *Principal 4* and the *exceptional reasons* clause in the ESRC Research Funding Guide[10].
2. Data access:
    a. data: not applicable as data is third party generated so data access provisions are not mandated (see above).
    b. Metadata:
        i. When data is used to support research findings (becoming *research data* under EPSRC definitions[11]) then metadata, as defined by the EPSRC[12], will be published in line with the policy detailed in section *A. Data collection and data access*. This is again inline with RCUK's *Principal 4.*
        ii. Additionally, meeting EPSRC Expectation II and RCUK's *Principal 4*, published research papers will include a short statement describing "how and on what terms any supporting research data may be accessed" noting that this data is covered by the clause in the provided clarification document stating that access is restricted due to "compelling legal or ethical reasons [that] exist to protect access to the data"[13]. For confidential and/or sensitive data metadata will be limited to a researcher contact and a statement indicating that data access was under legal agreement and that "compelling legal or ethical reasons exist to protect access to the data".
        iii. Since these datasets are owned by a 3rd party and provided under agreement they do not fall under the categorization of an ESRC data asset[14,15] and therefore their meta-data will not be stored in the UK Data Service.

**Personal data collected by us**

1. Data collection: Currently no personal data is collected by N/LAB. If this situation changes then this document will be updated to include details of quality assurance during collection, digitisation/data entry, checking and additionally detailing processes to verify data authenticity in line with the ESRC/UK data Service guidelines[16].
2. Data retention: Currently no personal data is collected by N/LAB. If this situation changes, then that data will be held for 10 years from date of last use/access by a third party [EPSRC Expectation VII].
    a. See section A. Data Collection and data access.
3. Data access:
    a. data: will not publically accessible [EPSRC Expectation II, Expectation VI] and consistent with RCUK Common Principles on Data Policy Principal 4 and the exceptional reasons clause in the ESRC Research Funding Guide.
    b. meta-data: will be publicly published, including access statement detailing route to access [EPSRC Expectation V, Expectation VI]. If this data was collected under ESRC funding the

---

[8] https://www.epsrc.ac.uk/files/aboutus/standards/clarificationsofexpectationsresearchdatamanagement

[9] https://www.ukri.org/files/legacy/documents/rcukcommonprinciplesondatapolicy-pdf/

[10] page 12 (accessed 17th May 2018) https://esrc.ukri.org/files/funding/guidance-for-applicants/research-funding-guide/

[11] "Research data is defined as recorded factual material commonly retained by and accepted in the scientific community as necessary to validate research findings" https://www.epsrc.ac.uk/about/standards/researchdata/scope/

[12] structured metadata: "sufficient to allow others to understand what research data exists, why, when and how it was generated, and how to access it" https://www.epsrc.ac.uk/files/aboutus/standards/clarificationsofexpectationsresearchdatamanagement

[13] https://www.epsrc.ac.uk/files/aboutus/standards/clarificationsofexpectationsresearchdatamanagement

[14] https://esrc.ukri.org/funding/guidance-for-grant-holders/research-data-policy/

[15] https://esrc.ukri.org/files/funding/guidance-for-applicants/research-funding-guide/

[16] https://www.ukdataservice.ac.uk/manage-data/format/quality

meta-data will be provided to the UK Data Service. See section A. Data Collection and data access.

**Non-personal data collected by us**

1. Data collection: For each data set collected an additional document will be produced documenting the quality assurance procedures undertaken including details of quality assurance during collection, digitisation/data entry, checking and additionally detailing processes to verify data authenticity in line with the ESRC/UK data Service guidelines. These will be made available as part of the data collections documentation.
2. Data retention: 10 years from date of last use/access by a third party [EPSRC Expectation VII]
    a. See section A. Data Collection and data access.
3. Data access:
    a. data: data will be publically accessible, or if generated as justifiably commercially confidential not publically accessible [Expectation II, Expectation VI]. This will be made within See section A. Data Collection and data access. If this data was collected under ESRC funding the data will be provided to the UK Data Service within 3 months from the end of the grant as per the ESRC research data policy[17].
    b. metadata: data will be publicly published, including access statement detailing route to access [EPSRC Expectation V, Expectation VI]. See section A. Data Collection and data access. If the data was collected under ESRC funding the meta-data will be provided to the UK Data Service within 3 months from the end of the grant as per the ESRC research data policy.

## D. Data Security Summary: Main risks to data security and mitigation

See additional document *N/LAB Data Security Plan.*
All references refer to the relevant sections within the additional document.

1. Digital data corrupted/lost (hardware faults, onsite physical damage, user error, stolen equipment)
    a. Database is backed up to both a replica backup server (real-time backup file replay) and offsite tape in encrypted form (daily).
    (see *C. Data Backup*)
    b. Non-database data is backed up offsite in encrypted containers on offsite network drives.
    (see *C. Data Backup*)
2. Improper access
    a. remotely:
        i. All data stored and processed on specially configured computers .
        (see *A. Information Security: Data server, Data processing machines*)
        ii. All data is accessed only by designated users.
        (see *A. Information Security: Access management*)
        iii. All data access is logged and continuously monitored.
        (see *A. Information Security: Data server*)
        iv. All data access machines are securely configured.
        (see *A. Information Security: Data processing machines*)
        v. Strong passwords are enforced and two factor authentication used when required.
        (see *A. Information Security: Data server, Data processing machines*)
        vi. All data transferred offsite for backup is encrypted beforehand, preventing data access within the backup system. Backups are maintained within the University by Information Services under University policy.
        (see *C. Data Backup)*
    b. physically (i.e break-in, equipment stolen):

[17] https://esrc.ukri.org/funding/guidance-for-grant-holders/research-data-policy/

i. Data access computers and servers are fully encrypted at rest, forced or otherwise, rendering data inaccessible.
(see *A. Information Security: Data server, Data processing machines*)
ii. Password protected lock screens are used in all data access computers, protecting the data while machines are on but the operator is away.
(see *A. Information Security: Data server, Data processing machines*)
iii. Physical access is restricted and logged
(see *B. Physical Security*)

c. data interception during transfer
i. Data is only ever transferred over encrypted channels or,
(see *A. Information Security: Data transmission*)
ii. Data is encrypted before being placed on/in an unencrypted channel/device
(see *A. Information Security: Data transmission*)

## E. Responsibilities

1. A nominated member of the research team will be designated as the data protection officer (DPO) responsible for implementing the data management plan, overseeing information security and ensuring all new members of N/LAB are correctly inducted with regard to their responsibilities with respect to this data management plan. In addition the DPO is responsible for keeping a record of 3rd party data and any personal data collected by us including: source of the information, who we are able to share it with, the reason for having the data and how long it will be kept. The DPO will review these documents as required and at least once a year.
**Current N/LAB DPO:** Dr. Gavin Smith (gavin.smith@nottingham.ac.uk), N/LAB Data Science Lead. The N/LAB DPO is distinct and in addition to the institution appointed Data Protection Officer who is responsible for the University's (including N/LAB's) overall GDPR compliance. Full details, including the current GDPR DPO can be found on the University webpage[18].
2. For 3rd party data acquisition, responsibility is assumed, under the guidance of this data management plan and the designated data protection officer (DPO), by the research team member responsible for the relevant company contact.
3. In the case of data collection by us: All aspects except information security will be the responsibility of the researcher undertaking the data collection, under guidance with regard to the data management plan and the designated data protection officer (DPO).

## F. Resources

1. All resources required for current research have been acquired. Data capacity will be upgraded as necessary if this situation changes.

[18] https://www.nottingham.ac.uk/governance/records-and-information-management/gdpr-overview.aspx