Guidance for the preparation of data management plans for ground engineering projects

Edition 1 – March 2024

DRAFT FOR PUBLIC CONSULTATION

Published by
Association of Geotechnical and Geoenvironmental Specialists

AGS Association of Geotechnical & Geoenvironmental Specialists

Copyright

Association of Geotechnical and Geoenvironmental Specialists
Forum Court, Office 2FF,
Saphire House,
5 Jubilee Way,
Faversham,
Kent,
ME13 8GD,
United Kingdom

Tel. 020 8658 8212 Email. ags@ags.org.uk Website: www.ags.org.uk

Disclaimer

Although every effort has been made to check the accuracy of the information and validity of the guidance given in this document, neither the members of the Working Party nor the Association of Geotechnical and Geoenvironmental Specialists accept any responsibility for misstatements contained herein or misunderstandings arising here from.

The example plans given in this document are provided as a guide only. They are used to demonstrate considerations that should be made when writing a Data Management Plan, and can be used as a template as a starting point. The examples are not off-the-shelf solutions that are ready to be used, and must be customised to target your organisation's needs and requirements.

Contents

List of figures

List of tables

Acknowledgements

AGS Data Management Working Group

This document was prepared by the Association of Geotechnical and Geoenvironmental Specialists (AGS) Data Management Working Group (DMWG) with additional representatives from the following working groups:

- Contaminated Land Working Group;
- Geotechnical Working Group;
- Instrumentation and Monitoring Working Group;
- and Laboratories Working Group.

The AGS acknowledges the generous time and resources given by the individual members and their employers.

Main Authors

Craig Brown	BAM Ritchies
Jackie Bland	Structural Soils Limited, Chair AGS DMWG

Contributors

Claire Daly	Hydrock Consultants
Conrad Stewart	Harrison Group Environmental Limited
Dimitris Xirouchakis	Structural Soils Limited
Emma Cronin	SOCOTEC
Georgina Donbroski	RSK Geosciences
Ian Nicholson	Fugro Limited
Judith Nathanail	Land Quality Management Limited
Neil Chadwick	Digital Geotechnical Limited
Simon Miles	Atkins Réalis
Tom Smith	Ørsted
Tony Daly	Amageo Limited
Will Farden	Chemtech Environmental Limited

Foreword

Following on from the AGS Data Conference in December 2022, it was highlighted that there were many questions in relation to data management plans under BS 8574 – "Code of Practice for the Management of Geotechnical Data for Ground Engineering Projects". The AGS Data Management Working Group (DMWG) decided that the best way to address these questions was by producing these guidance notes on how to write an effective Data Management Plan. In June 2023 a subgroup of the DMWG was formed tasked with creating this document.

Good data management practice is key to the success of ground engineering projects, and it is hoped that this guidance will help drive better collection, processing, sharing and use of data throughout the industry.

Craig Brown
Chair of AGS DMWG Data Management Plans Subgroup

Part 1 – Guidance notes

1 Introduction

This guidance document aims to give practical advice in how to prepare an effective Data Management Plan (DMP) for use on ground engineering projects. Ground engineering covers a broad spectrum of disciplines with different requirements and types of data that are collected, stored and shared.

In March 2014, BS 8574 "Code of practice for the management of geotechnical data for ground engineering projects" was published to provide a framework for good data management practice in the industry. This standard sets out key requirements for a data management system, and highlights core principles that should be considered when planning for effective data management. BS 8574 covers all aspects of ground engineering. Therefore, it is beyond the scope of the standard to go into specific details of considerations that should be made, and as such the style and content of data management plans vary greatly across the industry.

By publishing this guidance, the AGS hope to improve data management practice across geotechnical engineering, and highlight how good data management is beneficial for a project or organisation's success.

This guidance provides a background on the core requirements of data management under BS 8574, discussing practical considerations that should be made for compliance. This document includes an example DMP, written from the point of view of a ground investigation, to demonstrate these considerations.

1.1 Terms and abbreviations

Table 1 provides definitions of terms and abbreviations that are used throughout this document.

Table 1 - Terms and abbreviations

Term or Abbreviation	Definition
AGS Data	Ground investigation data that has been compiled following the AGS Data Format specification (AGS, 2022)
BEP	BIM Execution Plan
BIM	Building Information Modelling, as per ISO 19650
BS 8574	BS 8574 Code of practice for the management of geotechnical data for ground engineering projects
CDE	Common Data Environment
Data	Facts and statistics, or information that has been compiled in a way that can be moved or processed
Data Format	Data that has been structured in a repeatable way with set rules and to a common specification
Data Management	Process of collecting, storing and sharing data

Term or Abbreviation	Definition
Data Management Plan, DMP	Written document outlining the process in which data is managed within an organisation or project
Data Manager	As per BS 8574 – "person responsible for the data management system"
Data Steward	As per BS 8574 – "person responsible for carrying out the process needed for the data management system"
Deliverable	A file or data item that is required to be delivered on a project
Information	The presentation of data in a useable format
Information Management Plan	A plan for the management of information
ISO 19650	International Standard – ISO 19650 Building Information Modelling
ISO 27001	International Standard – Information Security Management
Metadata / Attributes	Data that describes other data, can be used to tag and give context to data items
MIDP	BIM Master Information Delivery Plan
Organisation	A company or organisation
Validation	As per BS 8574 – "control technique used to detect data that are in the correct format and within acceptable limits"
Verification	As per BS 8574 – "act of checking transferred data… by comparing copies of the data before and after transfer"
Work Item	A subprocess for an activity within the main DMP process.

2 Data management framework under BS 8574

BS 8574 sets out a structure for documenting and planning for data management. It defines roles and responsibilities and breaks the framework down into defined documents with a clear hierarchy. The framework of BS 8574 is scalable to different organisational sizes and structures.

BS 8574 is primarily focused on logical data, which can be defined as structured data such as that stored in a database. It does not explicitly cover representation or evaluation of data which are topics generally covered by an organisations Information Management Plan. In practice, DMPs frequently expand upon the term data to include physical files (computer files such as documents or images), whilst adopting the principles of BS 8574. This guidance will explain how both aspects can be covered.

2.1 Roles and responsibilities

Two defined roles are given under BS 8574: Data Manager, and Data Steward.

2.1.1 Data Manager

A Data Manager is defined as an individual with overall responsibility for implementing and maintaining a data management system (BSI, 2014). This responsibility includes the preparation and maintenance of plans, reporting performance of such systems and the managing of Data Stewards by providing support.

If required due to the size and complexity of an organisation, there may be multiple Data Managers responsible for different parts of a data management system. This hierarchy should be documented within the system so that each scope is clear. For example, an organisation may employ a Data Manager to oversee the whole framework, but there may be departmental Data Managers responsible for specific work functions operating under different plans.

2.1.2 Data Steward

Under BS 8574 a Data Steward is defined as anyone carrying out processes or work under a data management system or plan (BSI, 2014). This effectively means anyone who is producing, transforming, using or sharing data in an organisation. A Data Steward is not a named individual, but an umbrella term used for many disciplines and role functions.

Responsibilities of a Data Steward are to handle data in line with plans and procedures, and to report back the effectiveness of these systems to the Data Manager.

In practice, the term Data Steward may be too broad for an organisation to use this term alone. Often additional roles are defined within an organisation to identify the responsibilities an individual has under a Data Management Plan. An example of this would be the employment of a Data Controller to manage sharing of data, or a Data Engineer responsible for processing of data into a system. This organisational structure is not set by standards because it is dependent on the size and complexity of an organisation, and the needs for these roles should be identified by an organisation's Data Management Strategy.

2.2 Data management system

BS 8574 sets out that the process of data management should start with establishing a data management system formed of a policy, strategy, manual and plans. Figure 1 demonstrates the hierarchy these components form in a data management system, and the key differences of what the components contain is given in Table 2.

Figure 1 - Data management system hierarchy

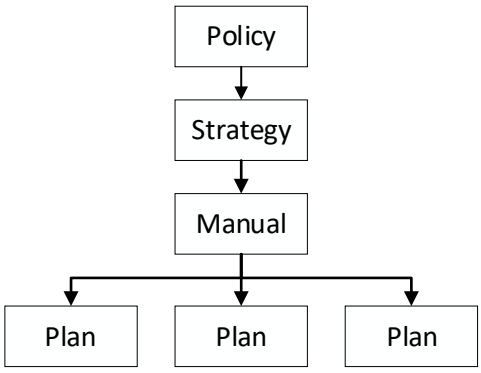


Table 2 - Key differences between policy, strategy, manual and plan

Document	Description and contents
Data Management Policy	• Give clear aims and responsibilities • Commit to continuous improvement • Be compatible with other policies covering quality and information technology • Creates a framework for data management
Data Management Strategy	• Set out goals such as Key Performance Indicators (KPIs) and Critical Success Factors (KCFs) for evaluating the data management framework • Set out timescales for achieving these • Effectively, the strategy sets out what the organisation wants to achieve, gives key measurable targets to assess if these goals have been achieved, and sets a blueprint of how to get there • Identify need for resources such as people, computer systems and software
Data Management Manual	• Plan for how organisation manages the data management strategy, using "Plan Do Check Act" (PCDA) method • Document the organisational structure in terms of data management • Bring together multiple data management plans in a register and show links • Gives responsibility for the data manager to review the data management system
Data Management Plan	• A plan of how data will be managed for a particular topic, this may be a department, project, or individual work item • Scope of plans shall be defined with clear boundaries • A large or complex topic may be split into multiple plans and brought together through references and by the Data Management Manual

The primary focus of this guidance is on Data Management Plans, and it does not go into detail about how to write effective policies, strategies or manuals.

2.2.1 Types of Data Management Plan

Data Management Plans may fall into two categories: organisational and project. Organisational DMPs will document how data is managed as a whole within an organisation, this may include data that is not linked directly to a project, or data that spans multiple projects. Project DMPs document how data is managed for a specific project or work item.

Plans will also differ depending on the type of organisation, and can broadly be generalised into client (or asset owner), consultant and contractor categories. The focus of the DMP may differ between these organisations due to their function on a project. For example:

- a client or asset owner may be focused on how they receive standardised data, maintain data, and store data for long term use;

- a consultant may have more focus on how they receive, interpret and transform data;
- and a contractor may be focused on how they produce, compile and share data.

2.3 Data Management Plan requirements

Data Management Plans prepared under BS 8574 should have a clearly defined aim and scope. Where multiple plans are employed in a system it should be clear how they relate to each other under the Data Management Manual.

BS 8574 defines activities undertaken as part of data management as work items. A work item may be defined as a particular task, or a subprocess within the wider data management process. An example of this may be that in the production of a report, laboratory data must be entered into a system, the entering of this laboratory data would be classed as a work item. A work item may also be a general stage in a data workflow, for example may relate to quality, storage, archive or disposal, these can be classed as a work item and referenced as a subprocess.

Large and complex data management plans may not be easily understood by the end user, especially if these work items are split between different departments or undertaken by different individuals. Although it is good practice for individuals to be aware of how their work affects the bigger picture, the purpose of a plan is to give a clear procedure for undertaking data management. For this reason BS 8574 states that plans may cover single or multiple work items depending on an organisation's requirements.

In practice, it is common that an organisation breaks down a data management framework into a plan for the organisation as a whole (covering general business processes), and project specific data management plans. It should be noted that depending on how an organisation is structured it may not be necessary to draft a bespoke plan for each project, it may be suitable to have a generic data management plan that is evaluated against each project to decide whether it is fit for purpose and meets the project requirements. This may be handled through an organisational project data management plan template, or by using appendixes for project specific details such as deliverable requirements and timelines, which can be completed on each project based on contractual requirements from a specification.

2.3.1 Content

Data Management Plans can be broadly broken down into the following topics:

- named roles and responsibilities;
- how data is collected, input and extracted from a system;
- how data is stored;
- what software or computer systems are used;
- what format is data transferred in (e.g. data/file formats);
- what media does sharing take place, including the timeliness or frequency of sharing;
- and data security requirements.

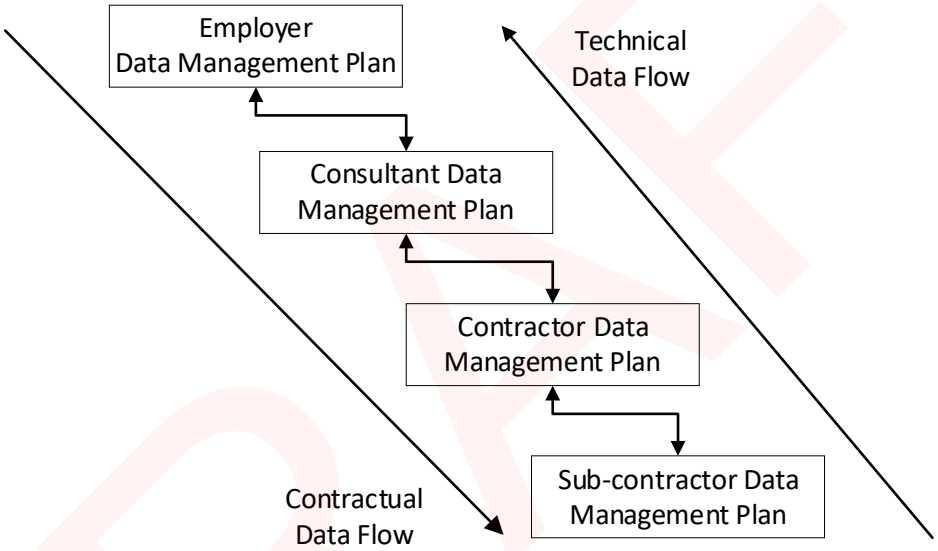
2.3.2 Aims and scope

Each Data Management Plan should state its objectives, and highlight the scope of which data or work items it covers. There should be plan(s) in place to cover all stages of a project within an organisation, and any interactions between the various stages.

It may be the case that a project as a whole has different Data Management Plans in place for each organisation in a supply chain. For example, an employer, a consultancy and a contractor may have their own data management plans, which link together at the point of sharing data. Each organisation will have its own computer systems and processes, and although there will be similarities on a project, a generic project level plan from one organisation is unlikely to be sufficient when passed to another organisation.

Figure 2 shows how Data Management Plans can link between organisations. In general if data passes directly between two organisations, their plans should align at the point of sharing. For this reason it is often the case that a Data Management Plan needs to be shared with the neighbouring parties, as the production of one organisation's project data management plan will inform the writing of the other. An example of where this occurs is if the sharing of data is to take place within an organisation's Common Data Environment (CDE).

Figure 2 - Example of Data Management Plan links with external organisations



2.3.3 Roles and responsibilities

Roles and responsibilities should be defined within a Data Management Plan. This outlines key responsibilities for individuals acting under the plan.

If particular roles fulfil actions of particular work items, these can be defined as additional organisation or project specific roles such as a Data Engineer, a Geotechnical Engineer, or a site operative. How this is broken down and how the roles are named are dependent on individual organisation's to identify.

2.3.4 Data systems and storage

Depending on project requirements data may be stored in multiple systems. For example, there may be a storage platform for informational documents, and a geotechnical database for storing raw data that aids the production of informational documents. As per BS 8574 data stored in multiple locations should be kept consistent.

BS8574 states that storage for data should be secure. Data security is a broad topic that covers multiple aspects including: prevention of unauthorised access, and prevention of data loss. This is discussed further in 2.3.7 Data security.

2.3.4.1 Data collection

BS 8574 states that data should be collected at point of source, and directly retained in a computer system. This is to reduce errors introduced through double handling or reprocessing. This good practice should be adopted wherever it is reasonably practicable. There are however times where it is not possible, for example where project security requirements prohibit the use of electronic devices on a site. If it is not possible to collect data digitally, a subprocess should be documented in the DMP of how this information is digitised. The digitisation of paper records should be undertaken as soon as reasonably possibly after collection to avoid any loss of data, and should be planned for with proforma templates so that there is control over the structure of how the analogue data is received.

Where practical to do so, data collection should utilise multiple choice lists to maintain consistency on a project. These lists may be defined by a data format or may be project specific. Multiple choice lists can be adopted on both digital forms through picklists, or from tick boxes on paper forms.

2.3.4.2 Unique referencing of data

There is a requirement under BS 8574 for data to be identifiable using a unique reference. This is applicable to data items within a database (such as a row of data relating to a sample), but could also be applicable for data files (referenced by a unique filename or metadata). A unique reference may also include any item that needs to be distinct within a project or organisation such as a location identifier or a laboratory schedule reference. Unique referencing must remain consistent for a data item to avoid accidental duplication.

Unique referencing may be formed of a combination of metadata. For example it may be formed of a location identifier, depth and sample type. In databases this is usually set through key fields configured in the database model.

Merging of datasets may cause issues with unique referencing, and this issue should be considered when designing the DMP. For example, merging of data may require additional validation checks, or cleansing of data through transformations. This merging may arise from addition of newly collected data, or from importing multiple large datasets such as multiple contractors working on a project.

File naming conventions will be dependent on the organisational or project requirements and the type of data being collected and transferred. It is good practice that file naming conventions are designed in a way that allows a file to be identifiable and distinct when placed outside of a folder structure.

2.3.4.3 Data validation and verification

A requirement of BS 8574 is that data verification and validation methods are used, and that these methods are repeatable, traceable and auditable. Validation and verification steps should take place at any point data is imported or exported from a computer system.

Definitions for validation and verification are given within BS 8574 as:

- validation is “an automated check to ensure the data are between acceptable limits and in the correct format”;
- and verification is an action “performed to ensure the data have been imported correctly”.

Validation steps may be programmatic checks undertaken by a computer system or manual checks undertaken by an individual. Programmatic validation steps may include data format structural integrity checkers, or data validation on mobile device logging systems checking if a value falls within a certain range (e.g. data field types, depth values, or text from a predetermined list). Validation checks that are unable to be programmed may be undertaken by an individual, this may be technical sense checks of data or checking paper records against defined rules.

Verification checks are often undertaken automatically within a computer system, such as a database confirming that all changes have been made without error. Verification steps can also be assigned to an individual, such as confirming that an import to a database has merged data in the way that was expected. An example of a manual verification check would be confirming that data has not been duplicated upon import into a database when merging datasets.

BS 8574 states that “newly collected data should be checked against existing data” for “quality and compatibility”. Quality is a broad term that may need definition in the DMP, or may be covered by an external document such as an organisations Quality Management System.

2.3.5 Data sharing and version control

BS 8574 states a requirement to cover the topics of timeliness, data status and version, exchange formats and sharing methods.

2.3.5.1 Timeliness

Timeliness sets out the timeline of how often a work item is undertaken or when the item is triggered. This frequently relates to the timeline of when data items are produced or shared, but may also relate to stages in a project lifecycle such as data archival or disposal. In this document deliverables are defined as a data item or file that is required to be delivered on a project; for example, this may be a drillers log, an AGS Data Format file or a photograph.

Listing the timeliness against individual work items may be complicated for end users of the DMP to digest, and so the example of this document introduces a concept of a “Deliverable Schedule”, demonstrating this information in a clear tabular form. An example of a deliverable schedule is provided within Appendix A – Deliverable Schedule.

2.3.5.2 Data status

Data status may relate to individual data fields, rows in a database (records), or data sets as a whole such as a data file. The status of data is important when sharing datasets as it can demonstrate the level of confidence in its contents. Example status definitions may be preliminary, draft and final. These terms may have different meanings between projects so should be defined in the DMP. Other document status may be used by a project or organisation and require definition, such as a release status defining the intended target audience of a file, such as tender, design or archive.

2.3.5.3 Revision versioning

File versioning is important to ensure consistency of data, so that up to date information is available to an organisation or project. BS 8574 gives a recommendation for data to be sequentially referenced for revisions, with a defined process for dealing with old versions.

File versioning may also be linked with user access rights to data so that approved versions are not accidentally modified in order to retain this consistency. Further details on access rights are discussed in 2.3.7 - Data security.

2.3.5.4 Sharing and data exchange

BS 8574 states that transfer methods should be agreed at the onset of a project. This is so that there is a clear and defined process for managing sharing and data exchange. Sharing of data may differ in the level of complexity depending on the organisation or project requirements, this may range from simple file transfers to fully integrated links between computer systems.

The use of industry standard exchange formats are preferred when sharing data to ensure compatibility between organisations and for the long term readability/accessibility of data. For example, proprietary file formats may be inaccessible in the future if software vendors discontinue products, and files containing spreadsheets may encompass formatting that is handled differently between versions of the software. Examples of industry standard file formats given in BS 8574 include the AGS Data Format (AGS, 2022) and the Federation of Piling Specialists (FPS) electronic pile schedule.

There may also be a requirement on a project to share data with third parties. An example of this may be conditions in some publicly funded ground investigation contracts to share data with the British Geological Survey (BGS) National Geoscience Data Centre (NGDC). Details of how and when this should be shared should be included in the DMP, as there may be caveats such as confidentiality periods before the data is exchanged. In general it may be good practice that written permission is sought from all data stakeholders to avoid legal issues when sharing data with third parties.

2.3.6 Data archive, retention and disposal

BS 8574 states that the DMP should cover the topics of data archival, retention and disposal.

Data archive refers to the storing of data for long term readability, and usually relates to a time period required for the retention of data. The data archive section of the DMP should cover the archive storage medium, and how to retrieve data from it.

The retention period for data within an organisation or project may differ depending on the type of data. Retention periods may be set by contractual, statutory or organisational requirements. As different files may have differing retention periods, the system of how this is managed requires documentation, and may inform the way other parts of the DMP are designed, such as including using metadata to tag documents.

The retention period of data may also need to consider how shared files are managed. For example, email communications may sit outside of a controlled project data store, but may still require disposal.

It may be the case that archive, retention and deletion procedures already exist within an organisation outside of the DMP. In this situation the documents should be referenced rather than duplicated, and it should be demonstrated where these procedures fit in the main data

workflow. It may also be the case that there are additional steps in preparing a project for archive or deletion before moving to these external procedures.

2.3.7 Data security

Data security is important to prevent the loss of data either through system failure, malicious attack or accidental data destruction. BS 8574 states that data should be regularly backed up. This is often undertaken automatically within modern computer systems. The process and timeliness of data backup should be documented in the DMP, or if these details are documented elsewhere in an organisation such as an IT policy, the external document should be referenced.

In addition to data backups, BS 8574 also states that there should be a disaster recovery plan in place. This includes details of how data is restored in the event of data loss. The complexity of this may be dependent on the size and complexity of a data system, and may be covered by existing policies within an organisation.

There is a requirement that data is safeguarded from unauthorised access. This may include safeguarding from both malicious attack from outside of an organisation, and from unauthorised access within an organisation. Access control is frequently implemented using user rights controlling whether a user has the ability to read or write data to a storage system. It may be the case that this is easier to define with user groups when planning for data management.

There is also an issue of data security related to sharing of data. The method of sharing chosen for a project or organisation should have a level of security adequate for the type of data being shared. For example, highly sensitive information may not be considered secure when transferred via email communications. Also details of how sensitive data such as that covered by non-disclosure agreements is shared is important so that these conditions are not accidentally breached.

2.4 Similarities between BS 8574 and other standards

There are possible overlaps in the aims of BS 8574 with other standards that may be adopted within an organisation or project. For example, both BS 8574 and Building Information Modelling (BIM) cover aspects of sharing data. The aims of these standards do not contrast, but one standard may not fully cover the requirements of the other so they often exist in unison.

Similarities with BIM and BS 8574 may include the requirement to plan for timeliness, unique referencing of data, detailing collection tools and specifying metadata. Within BIM these are frequently documented on a Master Information Delivery Plan (MIDP) as part of a BIM Execution Plan (BEP). The example DMP included in this guidance uses the concept of a Deliverable Schedule which has similar functions to a MIDP that can be used where BIM is not used on a project. An example of a Deliverable Schedule is provided in Appendix A – Deliverable Schedule.

Where aspects of data management are included in other policies or plans on a project they should be referenced and the information not duplicated to avoid any conflict.

3 Using this guidance

This guidance gives practical recommendations on how to write a DMP that complies with the data management practices outlined in BS 8574. Many of the aspects covered in this guidance are not explicitly mentioned in BS 8574 but can be inferred by the general principles it outlines.

When planning for data management, it may be beneficial to initially map the organisations current processes to understand how everything fits together before starting documentation. It may also be useful to keep things simple at first, and add complexity during the review and improvement cycle of the data management system.

There is no "one size fits all" approach to DMPs because they are governed by an organisation's internal processes, chosen software, and the type of data being collected. In addition, best practice is continuously evolving and so different approaches to this guidance can be taken.

3.1 Example Data Management Plan

Part 2 of this guidance contains an example data management plan given from the viewpoint of a ground investigation contractor. Although this example is provided for a particular viewpoint the requirements and guidance may be relevant to all types of DMP.

Headings used in the example are provided as a guide only, and may not be relevant to every organisation or project. This example is just one way of writing a data management plan and there are other approaches that can be taken. The example is arranged to demonstrate considerations that may be made when planning for data management to comply with the principles of BS 8574.

The example is not an off-the-shelf solution that can be directly implemented within an organisation or project; it requires customisation to the type of data that is being generated and specific needs and requirements of an organisation or project.

The example DMP provides requirements and guidance against each heading, and includes example text demonstrating these points. The guidance given in the example should not be considered a definitive list of everything that should be included, it is down to the author of the DMP to identify what is relevant to their organisation or project.

Part 2 - Example Project DMP

S1 Introduction

This Data Management Plan (DMP) covers the collection, processing, storage and transfer of geotechnical data as part of the *XYZ ground investigation* project.

The employer for the project is *XYZ*, the consultant is *XYZ* and the contractor is *XYZ*.

This DMP has been produced following the guidance of BS 8574.

Recommended content:

- Introduce project or organisation.

Guidance:

- This should give enough background information to provide context to the Data Management Plan. This should include main stakeholders and a brief description of the project or organisation.

S1.1 Scope

This DMP covers all geotechnical data collected throughout the ground investigation project and applies to all employees and subcontractors contributing to project data.

This DMP should be read in conjunction with the specification (*XXX-YYY-SP-01*), and in conjunction with the following documents:

XXX-YYY-ZZZ-01	Data Management Policy
XXX-YYY-ZZZ-02	Data Management Strategy
XXX-YYY-DMP-00	Data Management Manual

Recommended content:

- Should cover the aims and boundaries of the DMP, and highlight which stages of data/work items it covers.
- Should include references to other parts of the DMP framework such as the policy and strategy and how they link. This includes references to other DMPs and specifications (if applicable).

Guidance:

- This information can be provided with a table.
- May provide references to other documents such as quality management systems, data retention policies, Information Technology (IT) management plans, etc. if their content affects data management under this plan.
- A large project may be broken down into multiple DMP's in order to give clearer and more concise guidance to projects. For example, there may be a separate DMP for ground investigation than one produced for piling. These are linked through overarching documents.

S1.2 Governance

This document has been produced by the project's Data Engineer, and shall be reviewed at a period not exceeding six months.

Recommended content:

- Include details on when this document is to be reviewed, and who is responsible for it.

Guidance:

- This does not have to be a named individual and may reference a role in an organisation or project, such as a Data Manager.

S1.2.1 Reviewing

Reviews of this document shall take into account any findings from audits that have taken place, and any feedback provided by the client. Once reviewed this document shall be approved by the company Data Manager.

Recommended content:

- Include if applicable.

Guidance:

- If there are specific requirements on how this document is to be reviewed, they should be documented here.

S2 Definitions and abbreviations

This document contains the following definitions and abbreviations:

AGS Data	AGS Data Format file in accordance with AGS (2022).
BS 8574	BS 8574 Code of practice for the management of geotechnical data for ground engineering projects.
CDE	Common Data Environment
GDPR	General Data Protection Regulations
Validation	As defined in BS8574: "control technique used to detect data that are in the correct format and within acceptable limits".
Verification	As defined in BS8574: "The checking of transferred data, usually at the stage of input to a computer system, by comparing copies of the data before and after transfer"
Deliverable	Any item of data that is required under the contract to be delivered on a project.

Recommended content:

- Should include any abbreviations used in the DMP for quick reference.

Guidance:

- This should also include any key terms with specific meanings.

S3 Roles and responsibilities

Roles and responsibilities under this Data Management Plan are as follows:

Role Category	Position	Responsibilities
Data Manager	Senior Data Manager	Maintain data systems. Provide adequate data management training. Approve data management plans. Undertake audits.
	Data Manager	Review data management plan. Ensure role functions under this plan are being fulfilled.
Data Steward	*Data Engineer*	Prepare data management plan. Control and track incoming data, storage and processing of deliverables.
	Data Controller	Manage flow of incoming and outgoing data. Track status of transmitted information. Notify relevant individuals of incoming data. Ensure timely transmission of deliverables.
	Operatives	Collect data in line with this DMP.

Recommended content:

- This section defines the roles and responsibilities under this DMP. This should include the Data Manager and Data Stewards.
- Should list key tasks for each role.

Guidance:

- May also include other roles with specific responsibilities such as a *Data Controller*.
- A Data Steward under BS 8574 is anyone who collects, processes, stores or uses data. This umbrella term may be further split into specific functions. For example, a project may have a named role in charge of managing data, such as a *Data Engineer*.
- If this is a DMP specific to a project, named individuals may be included here.
- The involvement of these roles will be proportionate to the size and complexity of a project. They may be dedicated roles, or part-time roles undertaken as part of a wider role.

S3.1 Competencies and training

In line with the project specification, Data Managers shall have a minimum of 3 years of relevant experience in data management.

In line with the organisations policies, Data Engineers shall have a minimum of 3 years of experience in ground investigation. Data Engineers shall undertake training provided by the Data Manager for project data management prior to undertaking work on the project.

This section is only included if relevant to the organisation or project.

Recommended content:

- If applicable, should include details of what training is required and who requires it, in order to be able to fulfil responsibilities under this DMP.
- If applicable, should detail any restrictions on competencies for a specific role or function under this DMP.

Guidance:

- These competency or restriction requirements may originate from a project specification, company policies, or other specific requirement that has been identified when planning for the works. This may include items such as mandatory data security training, specific software training, or levels of experience.

S4 Systems and integrations

A diagram of core data systems used on this project, and their associated links, is shown in Figure 3. Links between systems where reasonably possible have been configured to transfer data directly, but some subprocesses require a manual transfer of files. Core systems used under this Data Management Plan are:

- data collection systems including:
  - company tablet data capture solution;
  - and data loggers;
- Software YYYY (geotechnical database);
- Software YYYY (Common Data Environment);
- various software packages including:
  - Software YYYY (reporting tools);
  - and Software YYYY (Computer Aided Design);
- and will interface with external platforms including:
  - Lab Platform YYY (laboratory information management system);
  - and Platform YYY (client Common Data Environment).

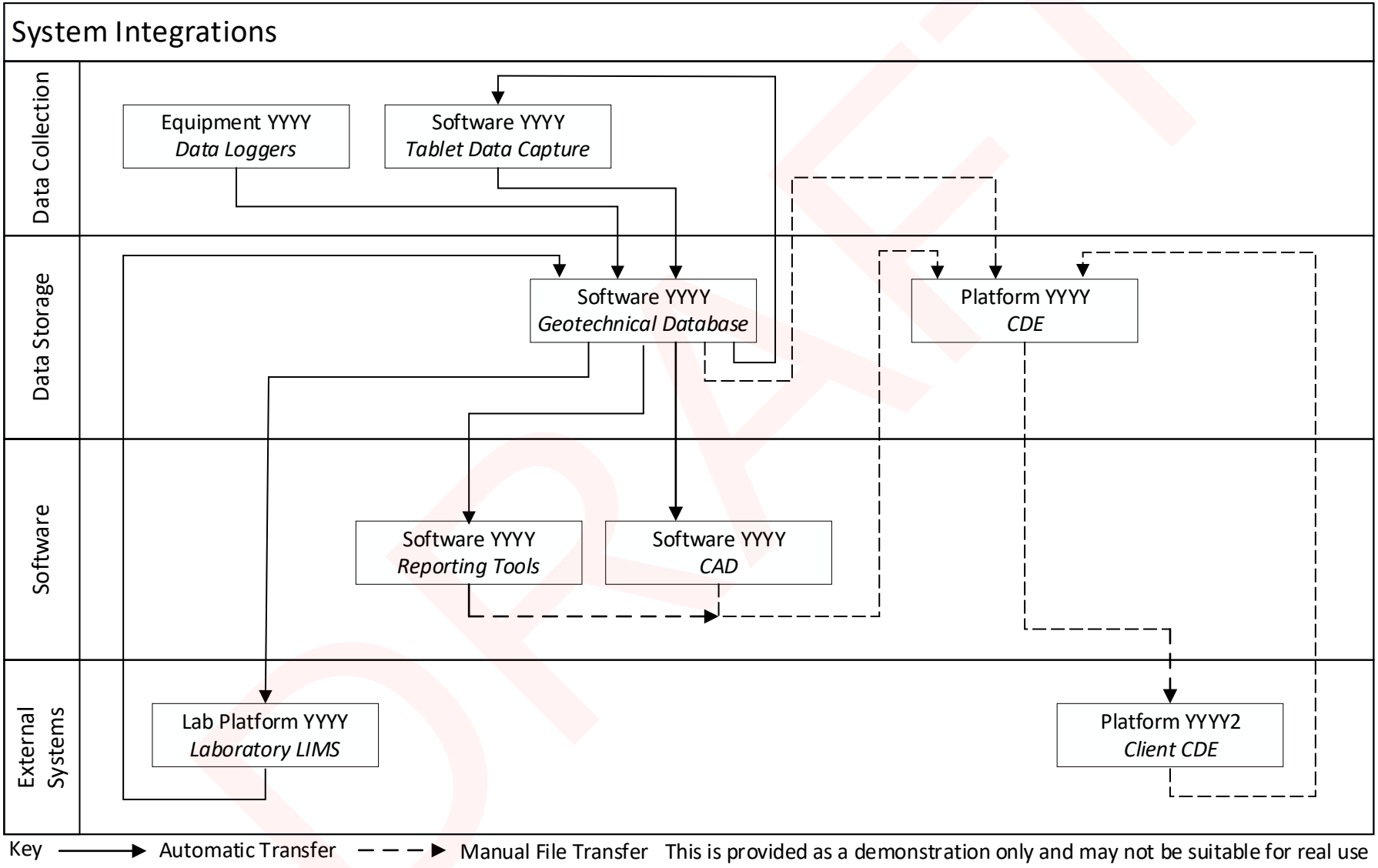
Recommended content:

- Should list key systems, software or data management platforms used under this DMP.
- Should include details of how these systems, software or platforms interrelate.

Guidance:

- This section can be split into subheadings for categories such as data collection and data storage. This may also be further subdivided into individual systems. The length and detail in this section will be proportional to the importance of the system, and size or complexity of the task it will be fulfilling.
- This section is not a detailed description of the process of how the systems are used, but should give high level details.
- This list does not have to be exhaustive of everything used on the project, but should focus on items that form key components in this DMP. An example of a key system would be the Common Data Environment (CDE) in use, or the primary data collection systems (e.g. a mobile device logging solution) used.
- A good way to present this information may be a simple diagrammatic map of links showing inputs and outputs from each component of a system.

Figure 3 - An example of a process map showing systems and integrations



S4.1 Data collection

All field data where reasonably practicable shall be collected digitally at point of source on specified digital forms relating to the work item. Where the use of digital forms is unavoidable due to site security restrictions, data shall be collected on approved paper forms, and digitised as soon as possible in line with procedures set out in section S7.2 - Paper to digital.

All software applications used for data collection, and the corresponding data collection forms used, should be kept up to date.

Field data collection

Field data will be collected through the company tablet logging solution. Forms have been configured to validate items against project data requirements and capture data in line with the AGS data format structure. The tablet logging solution pushes submitted data directly into the geotechnical database on the project.

Groundwater data loggers

The project requires the use of groundwater data loggers. The data loggers will be configured to transfer data automatically at set periods to the geotechnical database via telemetry. The YYYY data logging equipment has been chosen for data loggers on this project.

Recommended content:

- Should list data collection systems or manual data collection processes.

Guidance:

- Multiple data collection systems may be in use on a project depending on the type of data that is required. The data collection tool used for each work item should be documented in the DMP; a good way to display this information is a column on the Deliverable Schedule (see Appendix A – Deliverable Schedule).
- How data is merged from different collection systems should be detailed in section S7.4 - Database actions and merging of datasets.

S4.2 Data storage

There are two areas of data storage covered by this Data Management Plan: the Common Data Environment (CDE) and the Geotechnical Database. The CDE chosen for this project is Software YYYY.

All data should be stored on the CDE at the earliest opportunity, and documents worked on live from the CDE. The use of offline or local files should be avoided where reasonably practicable.

The client for the project also uses a CDE, with a "contractors area" for the sharing of data files. This system is Platform YYYY. Transfer of files to and from the clients CDE will be the responsibility of the project's *Data Controller* in the form of manual file transfers.

Recommended content:

- Should list data storage system(s) to be used, e.g. file servers or internal and/or external CDEs.
- Should detail any links between systems if applicable. For example, where there is an automated transfer of data from one system to another as part of the DMP.

Guidance:

- This section should not go into specific details such as file naming conventions or file metadata. Information regarding metadata and filenames should be included in S10 - Unique referencing of data.

S4.2.1 Geotechnical database

Geotechnical data is collated in the project Geotechnical Database through the use of Software YYYY.

This project has specific data requirements including customisations to the standard database. The database has been set up to use the "*AGS4_ProjectX*" database configuration model.

The database package has in-built reporting tools that are used to produce deliverables. These deliverables are to be stored within the project's CDE.

Recommended content:

- If applicable, any databases used to store data should be listed here.

Guidance:

- This may be a software package or platform that functions as a database.
- If the project requires a specific database model or configuration, the details of this should be given in this section.

S5 Data management overview

This DMP has been prepared following these principles:

- data should be collected digitally at point of source where reasonably practicable;
- data should be stored in a structured and identifiable way;
- and data should comply with the project specification.

An overview of the data management process is given in Figure 4. This process is further broken down into subprocesses for work items.

Recommended content:

- Should include high level details of how data is managed under this DMP.
- Should bring together detailed subprocesses documented later in this DMP.

Guidance:

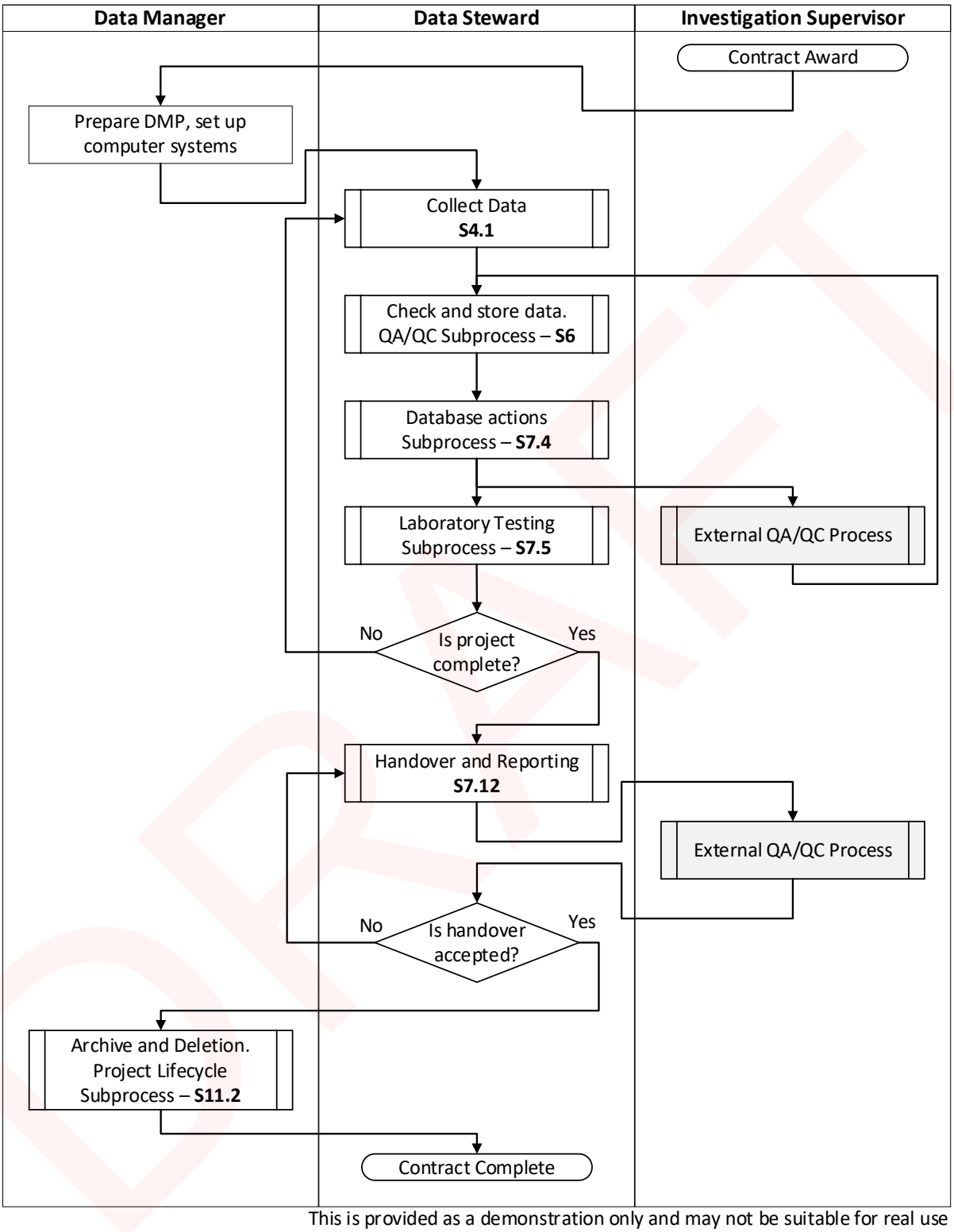
- The overview may be best communicated as a process map.
- This section is intended to be a high level overview of data on a project. This should be presented in a way that is easy to understand, such as visually demonstrated as a swimlane diagram (also known as a cross-functional diagram) of the main data management process.
- To make this DMP easy to use, it could reference further details of subprocesses documented later in the DMP, for example in sections:

  S6 - Quality assurance and quality control

  S7 - Processing of data

  S9 - Sharing and transfer of data
- Many data management workflows will include a mix of automatic and manual links between systems. It is recommended to make it clear whether steps are automated or require input from an individual. This may be by using different arrow styles representing the connections on a process map.

Figure 4 - Example of a data management overview swimlane diagram



This is provided as a demonstration only and may not be suitable for real use

S6 Quality assurance and quality control

Data items shall conform to the organisation's *Quality Management Plan* (XXX-YYY-ZZZ-19). The quality assurance and data storage process map is presented in Figure 5.

All data items shall be subject to validation and verification processes.

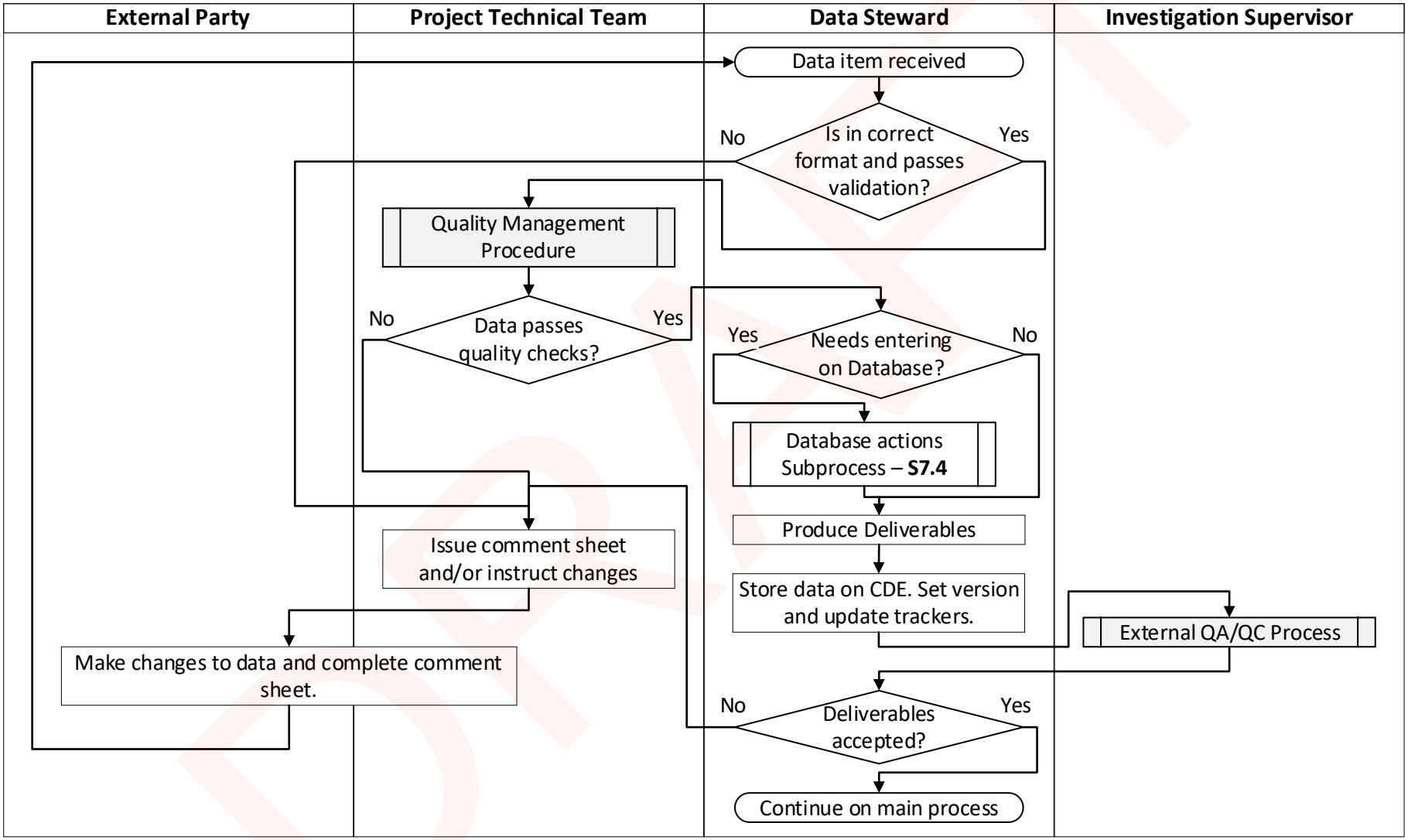
Recommended content:

- Should include details on compliance with applicable quality management systems in terms of the data workflow.

Guidance:

- The requirements of this section may be split into subsections depending on the complexity of the organisation or project.
- Quality processes may be split by general statements, or specific deliverable items (e.g. for a particular data file type).
- Where details of quality checks are documented in other management plans (e.g. quality management plans), these documents can be referenced here. Information should not be duplicated from other plans unless it forms a critical part of the data workflow process.

Figure 5 - Example of a quality assurance and data storage process map



This is provided as a demonstration only and may not be suitable for real use

S6.1 Validation and verification

Validation takes place at multiple stages of a data items lifecycle.

Data validation is used within the company's *tablet logging solution* via preprogrammed rules to check that data is within acceptable limits and key information is not omitted.

Each data item is quality checked by the project's *Technical Team* in line with the organisation's *Quality Management Procedure*.

The geotechnical database is configured to verify imports and exports during database actions. Imports into the geotechnical database shall be accompanied by a manual check by the Data Steward to confirm information has been entered into the correct place.

Recommended content:

- Should include details on the method of checking used, for example automated (preprogrammed) or manual (by an individual).

Guidance:

- Validation or verification may include manual checks undertaken by an individual. This may form a hold point in the data workflow and should therefore be documented here. A hold point is a part of the workflow process where an action is required before the process can be continued. For example, there may be an approval process within systems used (such as a CDE).
- Where this process is documented elsewhere it should be referenced instead of duplicating information, unless it forms a critical part of the data workflow process.
- When writing this section the specific meanings of validation and verification as defined in BS 8574 should be used.
- There are multiple elements to the term validation. It may refer to an automated check of data such as file format integrity checkers (computerised checks that data is in the correct structure), or it may also refer to human checks of technical data. Both are important validation steps and should be documented.
- An example of verification is an automatic check that a system received what was transferred, or a manual check that an import went into the correct places within a database.
- The level of complexity of this section will be proportional to the size and complexity of the project or organisation.

S6.2 Data audits

The project may be subject to a random data audit conducted by the organisation's Data Manager.

During an audit the following will be assessed: the CDE checked for compliance with metadata tagging of data files, timeliness of deliverables, and confirming that the correct data procedures have been followed. The geotechnical database will be audited on compliance with the AGS Data Format and against the project's specification.

Recommended content:

- If applicable, details should be recorded of who undertakes audits, and at what intervals.
- If applicable, brief details should be given of the purpose of audits, and what is audited.

Guidance:

- Audits may come in different types:
  system audits – has a system done what it is supposed to do;
  and data quality audits – have individuals handled data correctly.
- Audits can be an important part of the data management process as they can lead to improving DMPs during their review cycle, and identifying any weaknesses in procedures.

S6.2.1 Issues of non-compliance

Findings from the audit will be fed back into the improvement lifecycle of the DMP and the organisation's data systems by the organisation's Data Manager.

An audit report will be provided to the Data Steward of corrective actions and timescales required.

Recommended content:

- Where audits are required, details of how issues raised and handled should be documented.

Guidance:

- This may include details of how issues arising from data management are fed back into the continuous improvement of the data management systems.

S6.3 Status of data

Document status designations are defined by the specification and are summarised in Table 3.

Table 3 - Example of document status designations

Data status designation	Description
Preliminary	Partial data, such as an incomplete exploratory log
Draft	Data for approval
Final	Complete, checked and approved data
Interim	Special report status for interim reports of partial datasets (e.g. a subset of locations or monitoring dates)

Recommended content:

- Defines what data status mean, e.g. preliminary, draft, final and interim.

Guidance:

- Where data status is contained within metadata such as within a CDE, details of any status codes should be documented here. This is not the same as data file version numbers.
- This section is aimed at dealing with terminology used for status. Subprocesses covering document versioning are included in S9.4 - Version control.

S7 Processing of data

-

Recommended content:

- Include any subprocesses for processing data.

Guidance:

- This section covers subprocesses linking back to the high level process map.
- Include any important parts of the data flow that follow a set sequence, such as transformation of data, so that it can be repeated.
- Subprocesses may include, for example, data cleansing, lab scheduling, or daily deliverables which may have a unique approval process.

S7.1 Logging

Logging of exploratory holes shall take place on the organisation's tablet logging solution. Logging systems have been configured to use picklist options where applicable in line with AGS Data Format abbreviations.

Daily logs including drilling records and engineers field logs are subject to automatic validation on the logging system, and an additional technical check upon submission of data prior to inclusion in the geotechnical database.

Due to the timescale for laboratory testing required under the specification, daily logs shall initially be submitted via email to the Investigation Supervisor for approval ahead of being uploaded to the clients CDE.

This section is only included if relevant to the organisation or project.

Recommended content:

- Include if relevant.

Guidance:

- Record details on how logs are completed and any associated data notes.
- Include hold points for validation checks.

S7.2 Paper to digital

Paper records may be required in areas of the project that have restrictions on the use of electronic devices. Paper records shall only be used where electronic collection of data is prohibited.

Paper records shall be collected using the proforma data collection templates given in Table 4. These templates are provided in *Appendix XXX*.

Table 4 - Example list of proforma paper data collection templates

Deliverable Item	Template
Drillers Log	XXX-YYY-FRM-01
Engineers Field Log	XXX-YYY-FRM-02
Chain of Custody	XXX-YYY-FRM-03

Paper records should be processed into digital systems as soon as possible after collection, this generally should happen within one working day. Digitisation of paper records shall take place by using the company's tablet logging solution so that data is standardised on the project.

Recommended content:

- If paper records are to be collected on a project, the data collection process such as the forms that are used should be documented.
- Explain how data is transferred into digital systems.

Guidance:

- It is recommended that all data is collected at point of source, however it is understood there are rare cases where paper records are unavoidable. Where paper records are used, data should be collected on proforma templates to make sure data items are not missed.
- To avoid loss of data, paper records should be digitised as soon as reasonably possible after data collection. It may be beneficial where practicable for the original collector of the information to undertake the digitisation as they have the most knowledge of the data.

S7.3 Deliverable Intervals

Deliverable timelines are presented in Appendix A – Deliverable Schedule.

This section is only included if relevant to the organisation or project.

Recommended content:

- Include if relevant.

Guidance:

- If deliverables have a specific process in the data management workflow they should be documented here. For example, if daily deliverables such as drillers logs have an expedited initial approval process to allow for scheduling within strict timeframes.

S7.4 Database actions and merging of datasets

Files imported into the geotechnical database shall follow the procedure set out in Figure 6. All AGS Data Format files should be validated for structural errors prior to database import, and sense checked as to what information they contain (e.g. check that location identifiers are correct). When importing data into the database a list of changes should be reviewed for how many new items and updates are taking place to sense check the data is correct before import.

After changes are made to the database, manual checks should be undertaken to make sure the data has been merged correctly and is free from error in line with the quality procedures.

Recommended content:

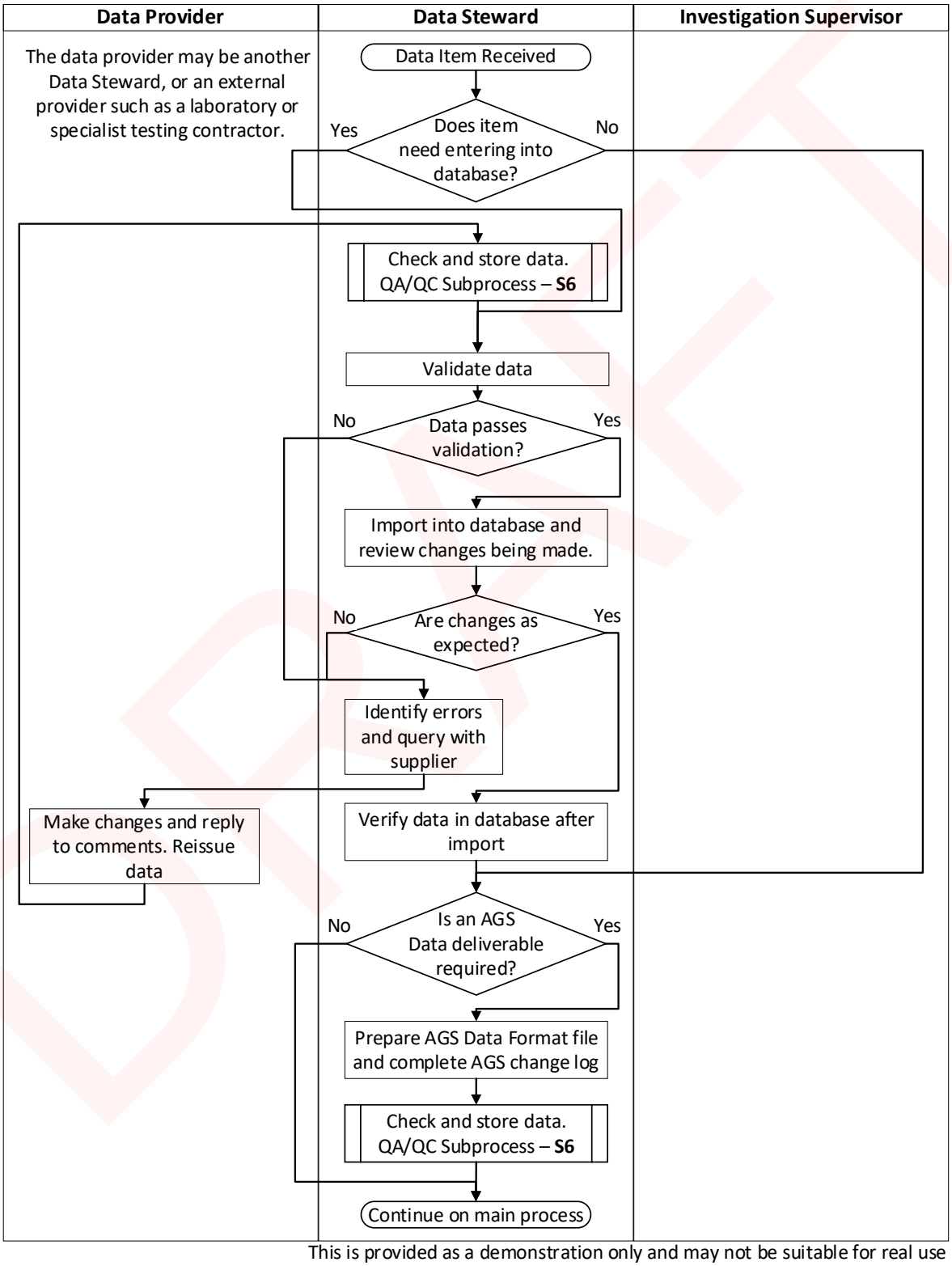
- Should provide subprocess for handling data merged from two sources into one database.
- Should give details on any checks that take place when inserting to or updating the database, such as checking key fields match etc.
- Should explain the process of verifying an import has worked, this may include manual checks.

Guidance:

- This section may consider how data from digital collection tools such as mobile device logging get merged into a live database.
- Common data merging issues that this subprocess aims to alleviate are: duplication of data where key fields have changed, accidental overwriting of newer information in a database with outdated data provided in a data file, and incorrect incoming information merged into a database from digital data capture solutions.
- Knowing the status of data being merged is important. If this is superseded data, how does that replacement happen? For example, receiving updated Cone Penetration Testing (CPT) data that needs to be merged with other drilling methods.
- If there are set rules for modifying data during a merging process they should be documented here as part of the subprocess. For example, this may be reconciling abbreviation codes in an AGS file or translating coordinate systems etc.
- Another consideration while preparing this DMP is how much of the data would need replacing on each data drop. For example, is the data completely replaced, i.e. does it get deleted before import, or are any tables kept such as sampling or scheduling groups in AGS data?
- This process may tie in with the writing of the specification when choosing the timescale and frequency of data drops on a project. For example, when asking for a data drop of whole project AGS data, how will the organisation manage merging this into systems? It should only be specified if the organisation or project are actually able to use the data and have need for it.
- It is good practice to keep a record of imports on a project. This may be a register documenting: the date and time, the type of import (e.g. overwrite, or new data), which data file and version is being imported, and who has undertaken the import.

- For large bespoke databases merging large datasets, it may be beneficial to use a "quarantine" or "sandbox" area to test the import prior to working on a live database.
- In some situations it may be wise to mandate the backing up of databases prior to merging.

Figure 6 - Example of database actions subprocess



This is provided as a demonstration only and may not be suitable for real use

S7.5 Laboratory scheduling

The process for undertaking laboratory scheduling is outlined in Figure 7.

Blank schedules are prepared with data taken from the geotechnical database and shall use the AGS data format key fields for sample references. Blank schedules shall be prepared using the template "YYY-BLANKSCHEDULE" and submitted to the Investigation Supervisor within the timescale provided in Appendix A – Deliverable Schedule.

Returned schedules shall be filed on the CDE and entered back into the geotechnical database. Laboratory schedules should be forwarded to the testing laboratory as both an Excel Spreadsheet and a digital AGS Data Format file containing the testing instructions.

The status of laboratory test schedules should be tracked using a laboratory testing tracker, and the geotechnical database should be kept up to date with testing restrictions and completion status of testing.

The laboratory testing tracker shall record: the unique schedule reference number, date sampled, date scheduled, date collected, date reported, shipping consignment number, testing laboratory, laboratory batch number (SDG ID) and laboratory report number.

Test definitions used on the project shall be as defined in Table 5.

Table 5 - Example of laboratory test name definitions

Suite E Soil Samples
Suite F Leachate Samples
Water Content
Liquid Limit, Plastic Limit and Plasticity Index
Particle Size Distribution by Wet Sieving
Sedimentation by Pipette

Laboratory results shall be provided by subcontractors in a single PDF file per reported schedule, and in AGS Data Format version 4.1.1. Reported results shall be imported into the geotechnical database.

Laboratories shall use the same sample references (AGS Data Format key fields) for reporting as they received via the schedule.

This section is only included if relevant to the organisation or project.

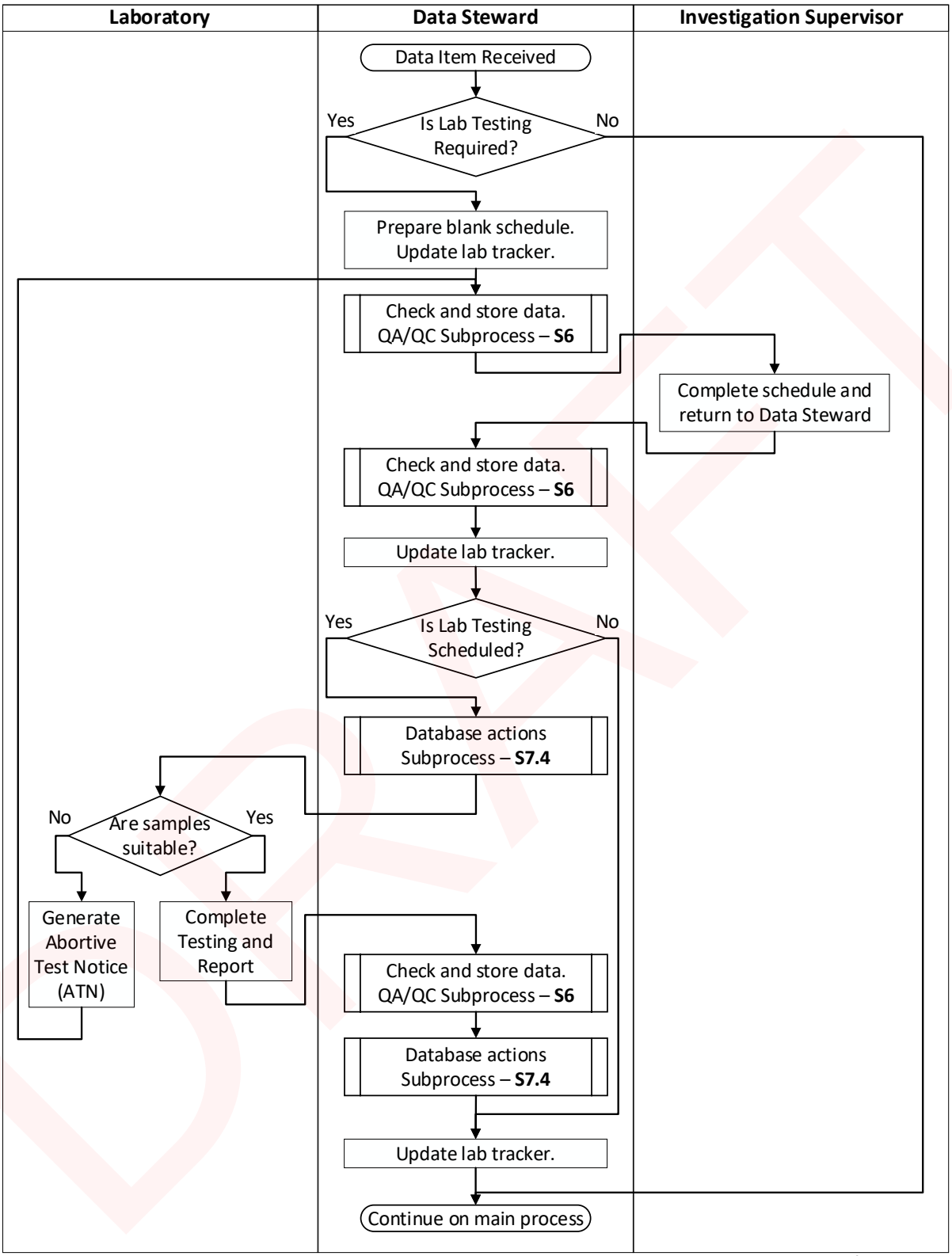
Recommended content:

- Should include details on how incoming and outgoing schedules are managed digitally.
- Should include any specific requirements for scheduling, this may include items such as test name definitions.
- Should document who prepares, completes, and submits schedules.
- Should include any applicable timescales on all parts of this process.

Guidance:

- If online portals are used to schedule testing or transmit schedule files, they should be documented here. Any tools used for scheduling such as bespoke templates should also be mentioned. This may refer to other documents detailing specifics of how a tool is used, such as a manual or a separate plan.
- This may include details of how incoming laboratory test result data is checked against schedules and how this test status is updated. For example, AGS scheduling information may need maintaining in the live geotechnical database.
- When writing this section consideration should be given to the management of data key fields (e.g. sample references), and what information is transferred at what timescale.
- This section should consider requirements for the organisation and project. For example, what data needs to be transferred, is there a requirement for it to be formed as a single exploratory hole per laboratory schedule for reporting requirements?
- This section should be written in conjunction with conversations with all parts of the supply chain to ensure that these requirements are feasible and compatible. This may only be applicable to a project DMP.
- It is good practice to detail how laboratory testing is tracked. An example of how this could be tracked is using the AGS Data Format laboratory scheduling groups (LBSG/LBST groups), but details of how to keep this data accurate and up to date is needed in the subprocess to make tracking effective.

Figure 7 - Example of laboratory scheduling process



This is provided as a demonstration only and may not be suitable for real use

S7.6 Monitoring

Monitoring data shall be collected digitally using the organisation's tablet logging solution. Monitoring points shall be identified by the AGS Data Format key fields for the monitoring point group (MONG). Monitoring data is transmitted digitally into the geotechnical database.

Monitoring points on the project are fitted with a monitoring datum marker that has an offset to the surveyed location datum. Corrections to the location datum are calculated by the entry sheet based off this value. All monitoring depths/distances shall be given in relation to the locations coordinates and the AGS Data Format field "Ground level relative to datum" (LOCA_GL) as stored in the geotechnical database.

This section is only included if relevant to the organisation or project.

Recommended content:

- Should detail how monitoring data is collected, received, processed and transmitted.

Guidance:

- If monitoring data requires transformation to merge with other systems, these details should be documented here. An example would be monitoring data that requires conversion to the AGS Data Format for inclusion with wider project data.

S7.6.1 Data loggers

Groundwater data loggers and instrumentation shall be set up to record in units and timescales defined by the project specification. All monitoring instrumentation shall be set to start recording on the hour so that all instruments have readings triggered at comparable times to allow for barometric compensation of data.

Barometric compensation of water level loggers will take place prior to inclusion in the geotechnical database by comparing data to the closest located barometric pressure instrument.

Data loggers transmit data via telemetry into the *Software YYY* platform where they are compensated for barometric pressure before being pushed into the geotechnical database.

This section is only included if relevant to the organisation or project.

Recommended content:

- Should include details of how data from data loggers is managed and processed.

Guidance:

- If part of the data processing workflow includes any transformations or calculations, such as compensating water levels from barometric data loggers, the procedure for this should be documented.
- This may include details of any telemetry systems used, or manual transfers of data that take place.
- This is an example subheading of monitoring. Repeat this section as necessary for other monitoring subprocesses.

S7.7 Data for analysis

Not required on this project.

This section is only included if relevant to the organisation or project.

Recommended content:

- Should include details on any subprocesses documenting how data is managed for use in analysis.

Guidance:

- For example, this may include details or references to plans concerning the AGSi data format (AGS, 2024).

S7.7.1 Modelling (2D/3D/4D)

Not required on this project.

This section is only included if relevant to the organisation or project.

Recommended content:

- Should include details on any subprocesses documenting how models are generated from data.

Guidance:

- This section isn't intended to describe how to build a model, but is aimed at how data is used within models. This may be procedures of how data is cleansed or incorporated into other data sets.

S7.8 Historical information

Not required on this project.

This section is only included if relevant to the organisation or project.

Recommended content:

- Should include details of the subprocess of how historical data/information is used on a project.

Guidance:

- This may take into account considerations such as the renaming convention of location identifiers, verification of data, transformation of data (such as units), or the assignment or reconciliation of computer codes (e.g. geology stratum codes in the AGS data format).
- Verification of data may include manual checks of the quality of data which should be defined here.

S7.9 Drawing production

Drawings shall be produced in the *Software YYY* Computer Aided Design (CAD) version 2024 software package. All drawings shall use the organisation template "XXX-YYY-ZZZ-50".

All drawings shall conform to requirements set out in the organisation's Drawing Management Plan (XXX-YYY-ZZZ-20).

This section is only included if relevant to the organisation or project.

Recommended content:

- Should include details of the subprocess of how drawings are produced and managed.

Guidance:

- If this guidance exists in other documents, they should be referenced here instead of duplicating them.

S7.10 Geographic Information System (GIS)

Geographic Information Systems (GIS) shall be managed in accordance with the project's GIS Management Plan (XXX-YYY-ZZZ-21).

This section is only included if relevant to the organisation or project.

Recommended content:

- Should include details on the subprocess of how data is moved in and out of GIS systems.

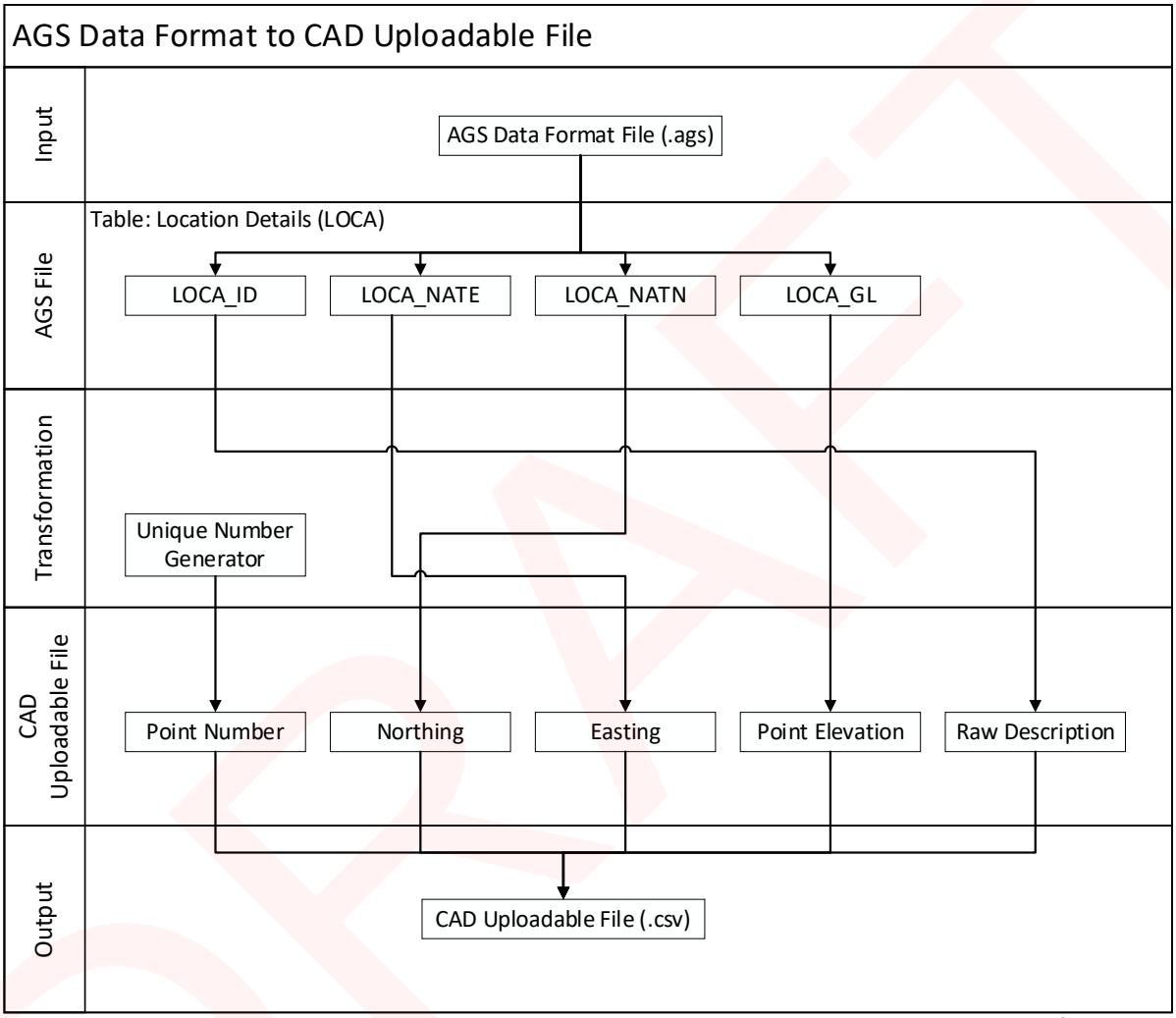
Guidance:

- If this guidance exists in other documents, they should be referenced here instead of duplicating them.

S7.11 Transformations and interpretations

Transformation of data is required between the AGS Data Format and the *Computer Aided Design (CAD) Uploadable File Format* during the production of site location plans. This process is undertaken through mapping of fields between the file formats and is demonstrated in Figure 8.

Figure 8 - Example of a transformation



This section is only included if relevant to the organisation or project.

Recommended content:

- If applicable, should include details of the subprocess of how data is transformed.

Guidance:

- An example of a data transformation may be how data is translated from one data format to another. This may also include transformations such as converting units in a received data file to make it compatible with other systems.
- This section will cover any Extract, Transform and Load (ETL) subprocesses. This may include manual steps, or automatic processes with details of how systems map between different formats.
- The example shown is for guidance only and is unlikely to be a real life transformation.
- If a transformation is a complicated process, such as an automatic process that maps database fields to a data format, this may be documented elsewhere and can be referenced in this section.
- If the conversion takes place as part of a standard function within software, it is not necessarily to fully document how this conversion works. In this situation a note of the software details used for the conversion is sufficient.

S7.12 Handover and reporting

Ground Investigation Reports (GIR) shall be compiled as a single bookmarked Portable Document Format (.PDF) file. As a minimum, bookmarks shall include: report headings, table and figure captions, and appendix titles.

Deliverable items included in the report shall include the most up to date version of documents. Reports shall be compiled from source documents for each revision to ensure that superseded documents are not included.

GIRs shall be checked in line with the subprocess S6 - Quality assurance and quality control.

GIRs shall be accompanied by a single AGS Data Format file generated from the same source of data.

Recommended content:

- Should document the subprocess of how reports or handover information is processed.

Guidance:

- This may include details such as how reports are compiled from information sources, such as any tools that automatically pull data from other sources on the CDE or database.
- This may include details of additional data such as the requirement for bookmarks in PDF documents.

S8 Deliverable data requirements

The file formats for deliverables is presented in Appendix A – Deliverable Schedule.

Recommended content:

- Should include details of deliverable types (i.e. file types for different deliverables), the timeline and/or frequency of deliverables, and any specific formats that data should be delivered in.
- Should give details on the approval process if applicable to the main data workflow.
- Should give details on how data is stored (e.g. folder locations).

Guidance:

- A good way to communicate this information is through a table format. This document refers to a Deliverable Schedule, see Appendix A – Deliverable Schedule. This is similar to the use of Master Information Delivery Plan (MIDP) within a Building Information Modelling (BIM) Execution Plan (BEP). If BIM is used and a MIDP is documented elsewhere fulfilling the purpose of this section, a reference to the document can be given.

S8.1 Specific data format requirements

Where reasonably practicable, all files shall be stored in established data formats with long term support, and proprietary formats should be avoided. Standard file types are provided in Table 6. Specific file types for each deliverable are documented in Appendix A – Deliverable Schedule.

Table 6 - Example of standard file types

Data Type	File Format	Description
Images / Photographs	.JPG	Joint Photographic Experts Group File (.JPEG)
Typed Documents	.DOCX	OpenOffice document format
Deliverable Documents	.PDF	Portable Document Format
Formatted Spreadsheets	.XLSX	OpenOffice document format
Unformatted Spreadsheets	.CSV	Comma Separated Values
Ground Investigation Data	.AGS	AGS Data Format
Drawings	.DWG	Drawing file
Spatial Data	.SHP, .GPKG	Shapefiles, OGC GeoPackage

Recommended content:

- Should detail what file formats are to be used.

Guidance:

- This section can give standard file formats for files, e.g. Portable Document Format (.PDF) for documents, Comma Separated Value (.CSV) for tabulated data, or that photographs should be stored as Joint Photographic Experts Group (.JPEG) files. It may also target specific data formats for deliverable items, such as AGS Data Format files (.AGS) for ground investigation data.
- This doesn't have to list a file type for each individual deliverable if they are covered by the generic statement.
- This section is not intended to include a specific description of how data is formed, only what data format the file is. Description of the data format structure itself is given later in section S12 - Specific dictionaries and data format structures.
- When choosing file formats, consideration should be given to long-term support of the data. This is often undertaken by choosing established open formats instead of proprietary file formats where reasonably possible.

S8.2 Handling of multi-file data sets

Deliverable items containing multiple files are to be included in a single unencrypted ZIP archive. File components within the archive shall be named so that they are identifiable.

Recommended content:

- Where a deliverable data item contains multiple component files, how these multi-file data sets are managed should be documented.

Guidance:

- An example of a multi-file data set may be where a data format submission is formed of multiple individual files, such as multiple Comma Separated Value (CSV) files that are linked.

S9 Sharing and transfer of data

The primary communication on the project is via email. Key email communications should be filed within correspondence areas of the CDE.

Deliverables should be shared electronically through the CDE.

Data shall be collected so that it can be shared within the timescale or frequency given in Appendix A – Deliverable Schedule.

Recommended content:

- Should include a high level description of the methods used to transfer data externally.

Guidance:

- This section should consider both incoming and outgoing data, and may cover data from multiple levels in a supply chain.
- Some example transfer methods are email, File Transfer Protocol (FTP), use of a Common Data Environment (CDE), or Application Programming Interfaces (APIs).

S9.1 Incoming data

Incoming data from the Investigation Supervisor will be downloaded and stored by the project Data Controller.

Incoming data from subcontractors will be uploaded to the CDE and a transmittal raised addressing the data to the project Data Controller.

Incoming data will be recorded on the project document register

Recommended content:

- Should detail the subprocess of how incoming data is managed.

Guidance:

- Details should be included about who data is transferred to (e.g. use of a shared mailbox or user roles on the project).
- If there are different processes for different parties they should be documented as subsections here.
- If Application Programming Interfaces (APIs) are used they should be considered here as part of the data management process.

S9.2 Outgoing data

Outgoing data to the Investigation Supervisor will be transferred to the clients CDE by the Data Controller.

Outgoing data will be recorded on the project document register.

Recommended content:

- Should detail the subprocess of how outgoing data is managed.

Guidance:

- Details should be included about who data is transferred to (e.g. use of a shared mailbox or user roles on the project).
- If there are different processes for different parties they should be documented as subsections here.
- If Application Programming Interfaces (APIs) are used they should be considered here as part of the data management process.

S9.3 Tracking of deliverables

A project document register will be maintained detailing all outgoing and incoming data transfers, this will be maintained by the Data Controller. This will record:

- file name and version;
- details on the sender and receiver;
- associated meta-data;
- date and time that the data was shared;
- and approval status of shared documents.

As part of the clients approval process, documents may be returned with comments. Comment sheets are recorded on the document register and notified by the Document Controller to the relevant members of the project team. Comment sheets will be stored on the CDE in the relevant area.

Tracking of the status of documents is undertaken through the document register.

Recommended content:

- Include details of how deliverables are tracked.

Guidance:

- This section should include any mandatory parts of the workflow that are used for tracking deliverables, this may be specific registers such as a document register etc. It should cover who completes and maintains this.
- This section should also cover how comment sheets are managed as part of an approval process; how they are stored, tracked and actioned.
- If this is tracked digitally in a CDE, the method and process should be documented here.

S9.3.1 File format change logs

AGS Data Format file deliverables shall be accompanied by a change log containing the following information:

- AGS Data File transmittal number (TRAN_ISNO);
- filename, version and issue date;
- purpose of transmittal (e.g. cumulative update / draft report / final report);
- remarks for use;
- and in compliance with the project specification, a list of sample key field modifications since last issue of AGS data.

This section is only included if relevant to the organisation or project.

Recommended content:

- Include if applicable.

Guidance:

- It is understood that in some circumstances change logs may be required for certain file types to assist in database management. An example would be change logs for AGS Data Format Files to assist in replacing data in a database. If change logs are required for AGS Data Format files, how they are recorded and what information is recorded should be listed here.
- The requirement of what information is needed will be different depending on the organisation or project. These requirements should not be too demanding, it is not feasible to request that all changes are documented, and will not be manageable by all parties involved both receiving or sending data.
- Good practice will be to only request the recording of changes that have been identified as critical changes in terms of the data workflow. An example would be if digital laboratory schedules are using the AGS Data Format, then any key field changes within the AGS Sample or lab testing groups require documentation. This is so that data is not duplicated when a database is updated.

S9.4 Version control

The version control of data on the CDE will be managed through its built in workflow functions.

Version control numbers follow the pattern outlined in ISO 19650-1 (BSI, 2019) and are demonstrated in Table 7. Version codes are defined with a published status (P or C code), followed by a major revision and minor revision number. Minor revisions are for internal revisions as part of changes made during the internal approval process

Table 7 - Example of data file version control codes

Status	Version code example	Comments
Work in progress	P01.01	Internal revision of a file, going through internal approval process
Shared	P01	Shared file (preliminary status), set after internal approval workflow is accepted
Published	C01	Published file (final status)

Recommended content:

- Should include details on the numbering process of data version control.

Guidance:

- The numbering process for change management may follow existing documented revisioning systems, such as that used in Building Information Modelling (BIM), e.g. P01.01, P01 and C01.
- The most simple file versioning system may be a sequential number for each revision.
- It is best practice to use an existing numbering system rather than inventing a new one.

S10 Unique referencing of data

All data on the project should be uniquely identifiable. Filenames should be unique and self-describing as set out in S10.1 - File naming and file metadata, and entries into the Geotechnical Database are identifiable by key fields in line with the AGS Data Format.

Recommended content:

- Should include general statement as to how data is uniquely referenced on a project.

Guidance:

- This may include details with a section on unique file naming, or how data entries destined for the geotechnical database are uniquely identified.
- Unique referencing is not confined to the examples given in this document. They include any work item that requires a unique reference, and may include laboratory schedule names, chain of custody records, etc.

S10.1 File naming and file metadata

All files on the project shall be uniquely named and identifiable by their filenames.

This filename is generated automatically based off the metadata fields configured in the CDE as shown in Table 8.

Table 8 - Example of table showing metadata stored about data files in the CDE

Metadata field name	Description	Example
Project	Project identifier	ABC123
Originator	Code representing originating company of data item	ACME
Volume	Volume the data relates to, such as the phase of the project. Default of ZZ (all volumes).	ZZ
Location	Exploratory location identifier	201
Type	Code representing the type of document	DL
Date	(Where required) Date that the data item refers to if produced as a daily deliverable, stored in YYYYMMDD format.	20240206
Depth Range	(Where required) Depth range associated with the data separated with underscore.	1.20_3.20
Role	Code representing the role on a project the organisation is taking. E.g. "W – Contractor".	W
Sequential Number	Sequential number given to documents automatically in the CDE. This distinguishes files containing similar metadata.	0001

Multiple choice options are used for applicable metadata, such as for document type and role. These are configured in the CDE and are documented in the organisation's Information Management Plan (IMP) XXX-YYY-ZZZ-04.

File names are generated within the CDE by joining metadata with a hyphen "-". Details of the metadata fields that are required for each deliverable type, and example filenames, are provided in Appendix A – Deliverable Schedule.

Recommended content:

- If applicable, should include details of any metadata that is required to be completed for deliverables.
- Should include details on how stored files are named.

Guidance:

- Use of metadata is not a requirement of BS 8574, but how files are uniquely named or labelled should be documented.
- An example of metadata is additional information that is attached to a file, such as a location identifier, or document type code.
- This may refer to a table in the DMP such as a Deliverable Schedule, an example is provided in Appendix A – Deliverable Schedule. Each deliverable should have its own entry on the Deliverable Schedule, this should detail any attributes that are required for each deliverable. This is similar to the use of Master Information Delivery Plan (MIDP) within a Building Information Modelling (BIM) Execution Plan (BEP). If BIM is used and a MIDP is documented elsewhere fulfilling the purpose of this section, a reference to the document can be given.
- It is not a requirement that Building Information Modelling (BIM) file naming conventions are used under BS 8574, but this may be a good starting point when planning the naming conventions. The BIM naming structure is likely to need adaptation with additional metadata to prove useful in management of geotechnical data.
- If this information is instructed in another document (e.g. an Information Delivery Plan), this document can be referenced here.

S10.2 Location naming

Exploratory hole location naming is chosen by the Investigation Supervisor and is as set out in the project specification.

Changes to location names in the event of redrills should be made in agreement with the Investigation Supervisor and will generally be formed of the original location name with a lowercase suffix starting with "a", for example "201a".

Locations should be given a new name if they have been moved after any data has been collected against the original location such as survey data.

Recommended content:

- Should include details of how location identifiers are formed and chosen.

Guidance:

- If there is a process of obtaining a location identifier from a system, this should be documented here.
- Details should be included on how to uniquely label additional locations such as relocated or revisited exploratory holes, for example by using a lower case suffix (i.e. 201a).
- With CDE's that implement BIM, consideration should be given into how location identifiers will affect the file naming convention. For example, in BIM some special character such as underscores and hyphens have a special meaning.
- It is inadvisable to put specific exploratory methods in location identifiers as this may change during an investigation. It is also inadvisable to use extensive metadata combined together to form a location ID.
- If the investigation is a multistage project, location identifiers should be unique between phases. For example, phase 1 may start "1XX", and phase 2 may start "2XX".
- Location identifiers should be kept as short and simple as possible to avoid confusion.
- When planning hole identifiers, consideration should be given to avoid conflicting information regarding a location. For example, an exploratory hole might be moved before it is drilled but survey data may already have been captured. With no alteration of location identifier, this may result in two survey reports for the same location.

S10.3 Coordinate systems

All geospatial coordinates on the project should be recorded in Ordnance Survey Great Britain 1936 (OSGB-36) coordinate system. All datums should be recorded from Ordnance Datum Newlyn (ODN).

Recommended content:

- Should include details of coordinate systems used on the project. This should remain consistent in data as far as reasonably practicable.

Guidance:

- Transformations may be required for project specific coordinate systems (local grids). If local grids, such as SnakeGrid are used, transformation details should be included here.
- If details of coordinate systems are established in other documents, they can be referenced here instead of duplicating the information.

S10.4 Sample referencing

Samples collected throughout the investigation will be assigned references at the point of collection in line with the AGS Data Format (AGS, 2022).

The AGS Data Format sample reference (SAMP_REF) field will increment sequentially for each sample taken. Duplicate environmental samples should contain the same key fields except for sample reference that continue to increment.

Geoenvironmental trip blank samples are required for quality assurance on the project. Where trip blank samples are taken they should be stored as a unique location with the prefix of "TRIP", e.g. "TRIP01". A register of these trip blank sample names and their associated sample batches shall be maintained.

All samples shall be assigned a unique identifier stored in the AGS Data Format sample ID field (SAMP_ID). For geotechnical samples this will be scanned from a preprinted barcode role to ensure uniqueness. For samples formed of multiple containers, such as geoenvironmental samples, a unique identifier is generated on the tablet data capture solution, with individual container barcodes stored in the AGS Data Format Sample Container Details Group (ECTN).

This section is only included if relevant to the organisation or project.

Recommended content:

- Should include details on how samples should be referenced.

Guidance:

- Referencing of samples should be kept consistent throughout a project. In general in ground investigation this is done by following guidance within the AGS Data Format (AGS, 2022).
- There may be unique situations where multiple samples are collected with similar references. For example, geoenvironmental sampling may require duplicate samples taken at the same depth. How these should be managed and identified should be documented, such as being distinguished by the AGS Sample Reference (SAMP_REF) field. This is so that each item maintains a unique reference.
- Geoenvironmental sampling may also require trip blank or field blank samples taken as part of a quality assurance process. If these are required it may require documentation on how the data is recorded, such as using a unique location identifier such as "FBS01" to mask the true location of a sample from a laboratory.
- The description of how references are used may be unique to the project or organisation. For example, the AGS Data Format sample reference (SAMP_REF) field may be a number that sequentially increments, or may be a sequential number at each unique depth and type.

S10.5 Laboratory and schedule referencing

Laboratory schedules will be labelled with the project identifier, a schedule type, an exploratory location identifier, and a sequential number. These components are separated with a hyphen. Locations shall be scheduled separately to maintain report order of results.

Geoenvironmental schedules will be marked with a schedule type of "ESH", and geotechnical laboratory schedules will be designated "GSH".

Examples of schedule names using this convention are:

- *ProjectID-ScheduleType-LocationID-SequentialNumber*
- ABC123-ESH-201-01
- ABC123-GSH-201-01

This section is only included if relevant to the organisation or project.

Recommended content:

- Should detail how laboratory schedules are labelled and identified.

Guidance:

- This may be dependent on project requirements, such as whether multiple locations can be scheduled together.

S11 Data and systems security

All software or platforms storing or transmitting data under this Data Management Plan shall be accredited to ISO 27001:2013, and approved for use by the Data Manager and organisation's IT department. All organisations acting under this Data Management Plan shall be accredited to ISO 27001:2013.

Access to all digital systems shall be controlled. Access to systems is maintained by the Data Manager using user credentials. Folders on the CDE are controlled by access groups, limiting areas to role functions. Details of folder restrictions on the CDE are documented in the organisation's *Information Management Plan (XXX-YYY-ZZZ-30)*.

Recommended content:

- Should include any specific data security requirements.

Guidance:

- If there are minimum IT security accreditations required they should be listed here. It should also be made clear if this restriction applies throughout the supply chain. The example shown uses ISO 27001 but other accreditations such as Cyber Essentials Scheme may be adopted, this is dependent on the organisation or project requirements.
- This section may include topics such as any geographic restrictions on where data can be stored, for example having to remain within the borders of a country.
- Details of how data remains secure during sharing should be included. This may be from defining access rights, with files or areas restricted to certain user groups.

- Consideration should be given to both internal digital systems, and external service providers such as cloud based platforms. Cloud based platforms may host the data within multiple data centres with different geographical locations (known as mirrors), if the geographic location of data is important, this should be recorded.
- When planning access restrictions it may be worth considering the "need-to-know" principle. This principle is that users should only be able to access information that they require to fulfil their work function. This is to minimise the risk of data leaks or inadvertent changes being made to data. An example of this may be to restrict GDPR sensitive documents or commercially sensitive information to certain user access groups.

S11.1 Backup policy

The CDE and geotechnical database are configured to maintain a backup of files in line with the *IT Data Security Policy (XXX-YYY-ZZZ-31)*. How data is recovered from this backup is included in the *IT Data Security Policy*.

Recommended content:

- Should include details on how data is backed up and restored in the event of a disaster (e.g. a disaster recovery plan).

Guidance:

- In some organisations this information may already be documented in an IT policy. Where it is documented elsewhere, these documents can be referenced here.
- Different systems may have different backup and recovery routines that may need documenting separately.

S11.2 Project lifecycle

The procedure for project lifecycle beyond the fieldwork and reporting period is shown in Figure 9. Upon report acceptance the project data will be subject to an archival and eventual deletion process.

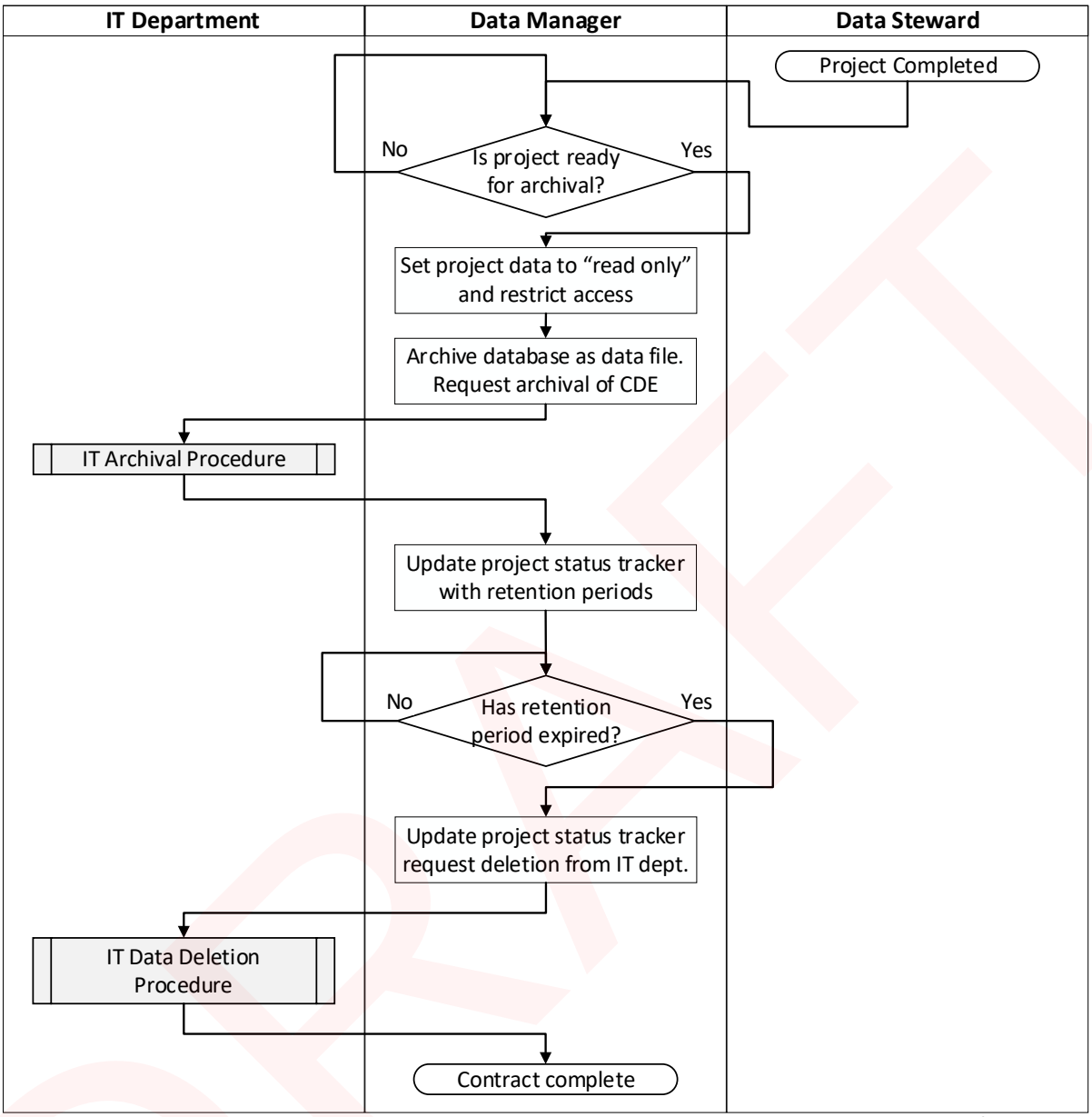
Recommended content:

- Should document what happens to data at the end of a project.

Guidance:

- This section should consider topics such as data retention periods, archiving data into long term storage, and data deletion.
- If there is a requirement to upload data to a third party repository such as the British Geological Survey (BGS) National Geoscience Data Centre (NGDC), this should be documented here. It is noted however that there may be implications on sharing data into the public realm, and that written permission should be agreed with all stakeholders of the data to avoid legal issues.

Figure 9 - Example of archival and deletion subprocess



This is provided as a demonstration only and may not be suitable for real use

S11.2.1 Archival

Upon confirmation of project completion, data will be archived.

When project data is due for archival, the geotechnical database will be backed up as a single AGS Data Format version 4.1.1 data file and stored alongside project data. The CDE will be marked as read-only and access rights limited to the Data Management team.

The moving of data to the archive area is administered by the IT department, and follows the process outlined in the *IT Data Archival (XXX-YYY-ZZZ-32)* procedure.

Recommended content:

- Should include details on the process of how data is archived for long term storage.
- Should detail when data is due for archival.

Guidance:

- If this information is documented elsewhere (e.g. in an IT policy), it can be referenced here rather than duplicated.

S11.2.2 Data retention and disposal

The data retention periods for the project shall be documented in the *project status tracker*.

Upon expiry of the data retention period, the Data Manager shall obtain confirmation of the intent to dispose data from the organisation's director. Upon confirmation the *project status tracker* is updated and the IT department notified that the data is ready for disposal.

The data disposal policy and further details on statutory data retention periods is governed by the *IT Data Disposal (XXX-YYY-ZZZ-33)* procedure.

Recommended content:

- Should include details on the process of how data is retained and deleted.

Guidance:

- Different types of data may have different retention periods that may be guided by company policy, project contractual requirements or legislation. The retention periods should be defined, and how data is managed during this period should be documented.
- If this information is documented elsewhere (e.g. in an IT policy), it can be referenced here rather than duplicated.
- Consideration should be made as to how data in different systems is disposed of. For example, using email for sharing data means that it will exist outside of the controlled storage ecosystem.

S12 Specific dictionaries and data format structures

-

This section is separated by each data format as subheadings. This guidance is applicable to all formats.

Recommended content:

- If applicable, any deviations from standard data format structures should be documented in this section.

Guidance:

- It is recommended that customisations to a data format are kept to a minimum to maintain the long term usability and compatibility of data. Consideration should be given to whether the information is actually needed, and if this information can already be stored in the existing structure.
- The size of this section and the level of detail will be proportional to the size and complexity of the organisation or project. In the event that there is extensive documentation for use of a data format it should either be referenced in an external document or included in an appendix to the DMP.
- Information on requirements relating to specific data formats should be included under subheadings of this section for each applicable format.
- Data formats often inform database design and the version specified, and any modifications, may lead to incompatibility between systems. Where a data format is to be used with databases, only one version of the format should be specified to ensure compatibility. Data format versions should not be changed part way through a contract to maintain consistency of data.
- The choice of data format and versions should take into consideration all users in the supply chain. It must be possible for the data to be produced as well as consumed. For example, new data format versions may not be implemented in software packages on day one of a version release and so would be unrealistic to specify.
- Any critical customisations to a data format should also be highlighted in the specification. Critical customisations are defined as changes which may involve modifications to a database design, such as adding fields and tables, or changing suggested data types.

S12.1 AGS Data Format

AGS Data Format version 4.1.1 (AGS, 2022) files are required on the project.

Cumulative AGS Data Format files shall be provided containing all available information for locations that have been fully logged at time of submission. Partial data for exploratory locations that are in progress shall be omitted from cumulative files. All AGS Data Format files shall be generated directly from the geotechnical database.

All AGS Data Format files shall be structurally checked using the AGS Data Format Validator.

The project specification requires that certain headings are populated with additional location information. These requirements are provided in Table 9.

Table 9 - Example of AGS Data Format heading guidance

AGS Group	AGS Field	Information Required
LOCA	LOCA_LOCA	Include name of location group as provided in Schedule 2 of the specification. E.g. “Substation 1”.

The project specification requires modifications to the AGS Data Format suggested data types as shown in Table 10.

Table 10 - Example of modifications to AGS Data Format suggested data types/units

AGS Group	AGS Field	Change
LOCA	LOCA_NATE	Data type: 3DP
LOCA	LOCA_NATN	Data type: 3DP
LOCA	LOCA_GL	Data type: 3DP

This section is only included if relevant to the organisation or project.

Recommended content:

- Should detail the specific version of the AGS Data Format that is required.
- Should document any requirements for the AGS Data Format that are not included in the standard AGS Data Format specification.

Guidance:

- Only one version of the AGS Data Format should be specified. Different versions of the format may have different units and field types which will inform database design and may not be compatible with each other.
- The version of the AGS Data Format chosen for a project should take into consideration all users in the supply chain; it must be able to be provided and consumed. For example, it may not be possible to do this on day one of a version release. The version used on a project should not change part way through to maintain consistency of data.
- Deviations from the AGS Data Format specification may be from AGS data field suggested types or units, custom group headings, or custom groups.

- If the organisation/project requires files to be attached to AGS Data Format files via the associated files group (FILE), this section should document how these files are formed and managed, and include details of what AGS Data Format group items they are tagged against.
- If there are critical customisations to the AGS Data Format, these should also be highlighted in the Ground Investigation Specification. Critical customisations are defined as changes which may involve modifications to a database design, such as adding headings and groups, or changing suggested units and types. Abbreviation codes are not considered a critical customisation.
- DMP(s) should be provided with the specification at tender stage if it is referenced by the Ground Investigation Specification. Some customisations to the AGS Data Format are only possible before data is collected on a project as they may inform a database structure.

S12.1.1 Custom groups and fields

Not required on the project.

This section is only included if relevant to the organisation or project.

Recommended content:

- If custom AGS Data Format groups are required, this section should include details of group structures.

Guidance:

- Should include at minimum: heading status, heading code, unit, data type, heading description and an example. It is recommended to use the same layout used for groups as used in the AGS Data Format specification (AGS, 2022) as this is a format that will be familiar to users.
- Any custom groups and fields should still abide by rules set out in the AGS Data Format specification (AGS, 2022).

S12.1.2 Abbreviation codes

Abbreviation codes for applicable AGS Data Format fields should be as per the AGS Data Format Abbreviations (ABBR) list available on the AGS website.

Project specific abbreviation codes are required as shown in Table 11.

Table 11 - Example of project specific AGS Data Format abbreviation codes

AGS Field	ABBR Code	Description
GEOL_GEOL	MGR	Made Ground
GEOL_GEOL	ALV	Alluvium
GEOL_GEOL	LC	London Clay Formation
GEOL_GEO2	S	Sand
FLSH_TYPE	Water	Water Flush
FLSH_TYPE	Polymer - X	Polymer X Flush

The AGS Data Format field "BGS Lexicon code" (GEOL_BGS) shall be populated using the "Computer Code" from the BGS Lexicon of named rock units (https://webapps.bgs.ac.uk/lexicon/).

Where a suitable abbreviation code is not available, agreement shall be made with the Investigation Supervisor for additional codes to be added to this plan.

This section is only included if relevant to the organisation or project.

Recommended content:

- Specify any AGS Data Format abbreviation (ABBR) codes for picklist fields, and their descriptions if they are not included on the standard AGS Data Format abbreviation set.

Guidance:

- It is not required to list codes that are part of the approved AGS Data Format abbreviation list, available on the AGS website: https://www.ags.org.uk/data-format/
- Project specific abbreviations should not conflict with existing codes on the published AGS Data Format abbreviation list.
- Some AGS Data Format headings do not have published standard abbreviations and will therefore always require an organisational/project specific set. These headings include:

  GEOL_GEOL Field geological descriptions > Geology code
  GEOL_GEO2 Field geological descriptions > Second geology code
  GEOL_BGS Field geological descriptions > BGS Lexicon code
  FLSH_TYPE Drilling flush details > Type of flush
  DISC_TYPE Discontinuity data > Type of discontinuity

S12.2 CSV formats

The project requires as-built location survey data to be delivered in a single cumulative Comma Separated Value (CSV) file with the headings shown in Table 12.

Table 12 - Example of CSV format details

Heading	Data Type	Description	Example
Point ID	Text	Location identifier	201
Easting	3 decimal place number	Coordinate easting	123456.789
Northing	3 decimal place number	Coordinate northing	987654.321
Height	3 decimal place number	Elevation datum	123.456
1D CQ	3 decimal place number	Height coordinate quality	0.013
2D CQ	3 decimal place number	2D coordinate quality	0.011
GDOP	1 decimal place number	Geometric dilution of precision	1.9
VDOP	1 decimal place number	Vertical dilution of precision	1.4
Date Time	Text date and time yyyy-mm-dd hh:mm:ss	Date and time of survey	2024-02-06 15:32:10

This section is only included if relevant to the organisation or project.

Recommended content:

- Should include details of headings and any data type requirements.

Guidance:

- This section is intended where CSV files are required for a data item in a specific format. For example, this could be for a certain layout of file that gets imported into specialist software, or to keep consistency between multiple contractors.
- CSV Format details given here do not refer to the AGS Data Format which should be listed under a separate section.

S12.3 Other data formats

-

This section is only included if relevant to the organisation or project.

Recommended content:

- Include any details that are required to ensure a data format file is completed to a required structure.

Guidance:

- Where data formats are unchanged from their original documentation the standard can just be referenced. Any deviations from the standard is documented here.
- This section is repeatable and should be renamed to the relevant data format being discussed.

S13 Standard forms and templates

Paper forms for data collection are attached in *Appendix XXX*.

Digital forms for data collection are shown in Table 13.

Table 13 - Example of a list of digital forms

Deliverable Item	Tablet Logging Software
Drillers Log	Daily Drillers Log v1
Engineers Field Log	Engineers Field Log v1
Monitoring	Groundwater and gas monitoring v1
Permeability Testing	Permeability Testing v1

This section is only included if relevant to the organisation or project.

Recommended content:

- Include a list of standard forms and templates that are required for data collection within the organisation or project.

Guidance:

- These may be included in an appendix and referenced in text.
- This may include copies of proforma data collection templates where data is restricted to non-digital collection methods.
- This is not intended to document what templates are used by a supplier, this section documents what templates are used in the context of this Data Management Plan.
- If standard forms are referenced in external documentation, such as within an Information Management Plan, it can be referenced here and not duplicated.

Bibliography and references

AGS (2022) *Electronic Transfer of Geotechnical and Geoenvironmental Data Edition 4.1.1*, Association of Geotechnical and Geoenvironmental Specialists. Available at: https://www.ags.org.uk/data-format/ags4-data-format/

AGS (2024) *AGSi Ground Model*, Association of Geotechnical and Geoenvironmental Specialists. Available at: https://www.ags.org.uk/data-format/agsi-ground-model/

British Geological Survey (BGS; 2024) *The BGS Lexicon of Named Rock Units*. Available at: https://webapps.bgs.ac.uk/lexicon/

British Standards Institution (2014) *BS 8574: Code of practice for the management of geotechnical data for ground engineering projects,* London: British Standards Institution. ISBN: 978-0-580-75023-6

British Standards Institution (2019) *BS EN ISO 19650-1: Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM). Information management using building information modelling – Concepts and principles*, London: British Standards Institution. ISBN: 978-0-580-92466-8

Institute of Civil Engineers, ICE (2022) *UK Specification for Ground Investigation*, London: ICE Publishing. ISBN: 978-0-7277-6523-9

Appendix A – Deliverable Schedule

Appendix A – Deliverable Schedule

An example of a Deliverable Schedule

Deliverable	Data collected in	Common Data Environment Metadata fields									Storage Location	File Name Example	File Type	Timeline	Comments
		Project Code	Originator Code	Volume	Location	Type	Date	Depth Range	Role	Seq. Number					
Drillers Log	Tablet Logging App	[PROJID]	ACME	ZZ	[LOCAID]	DL	[YYYYMMDD]	X	W	0001	/Data/DrillersLogs	ABC123-ACME-ZZ-BH201-DL-20240206-X-W-0001	PDF	Daily	
Engineers Field Log	Tablet Logging App	[PROJID]	ACME	ZZ	[LOCAID]	EFL	[YYYYMMDD]	X	W	0001	/Data/EngineersLogs	ABC123-ACME-ZZ-BH201-EFL-20240206-X-W-0001	PDF	Daily	
Core Photograph	Digital Camera	[PROJID]	ACME	ZZ	[LOCAID]	CPH	X	[TOP]_[BASE]	W	0001	/Data/Photo/CorePhoto	ABC123-ACME-ZZ-BH201-CPH-X-1.20_3.20-W-0001	ZIP (JPG)	5 days of location completion	Individual filenames must be able to identify location and depth.
AGS File	Geotechnical Database Software	[PROJID]	ACME	ZZ	X	AGS	X	X	W	0001	/Data/AGSFiles	ABC123-ACME-ZZ-X-AGS-X-X-W-0001	AGS	Cumulative, weekly and with reports	
Environmental Lab Schedule	Geotechnical Database Software	[PROJID]	ACME	ZZ	[LOCAID]	ESH	X	X	W	0001	/Data/Lab/Schedules	ABC123-ACME-ZZ-BH201-ESH-X-X-W-0001	XLSX	Daily (environmental)	
Geotechnical Lab Schedule	Geotechnical Database Software	[PROJID]	ACME	ZZ	[LOCAID]	GSH	X	X	W	0001	/Data/Lab/Schedules	ABC123-ACME-ZZ-BH201-GSH-X-X-W-0001	XLSX	5 days of location completion (geotechnical)	
Lab Report	External Platform	[PROJID]	ACME	ZZ	[LOCAID]	LR	X	X	W	0001	/Data/Lab/Reports	ABC123-ACME-ZZ-BH201-LR-X-X-W-0001	PDF	Upon completion of lab testing	
Exploratory Log	Geotechnical Database Software	[PROJID]	ACME	ZZ	[LOCAID]	EHL	X	X	W	0001	/Data/ExploratoryLogs	ABC123-ACME-ZZ-BH201-EHL-X-X-W-0001	PDF	5 days of location completion	
Ground Investigation Report	-	[PROJID]	ACME	ZZ	X	REP	X	X	W	0001	/Data/GIR	ABC123-ACME-ZZ-X-REP-X-X-W-0001	PDF	4 weeks of fieldwork completion (draft), 1 week of draft comments (final)	

This example is incomplete and requires populating based off organisational or project requirements.

Appendix A – Deliverable Schedule

Deliverable	Data collected in	Common Data Environment Metadata fields									Storage Location	File Name Example	File Type	Timeline	Comments

Feedback

Feedback is important to the ongoing development of this guidance, and it is understood that in the fast paced world of data management, good data management practices are continuously evolving.

Readers are invited to email the Association of Geotechnical and Geoenvironmental Engineers with recommendations of changes to this document.

Contact	**Subject**
ags@ags.org.uk	"AGS DMWG – Data Management Plan Guidance Recommendations"

Please address your email to the AGS Data Management Working Group and leave your name and organisation. To help organise changes we request you detail the relevant section number against your concerns or recommendations.

The AGS thank you for your involvement in developing this guidance.