



Dig Digital.

Work Digital. Think Archive. Create Access.

A guide to managing digital data generated from archaeological investigations

Infosheet #6 – selection strategies and digital data

As with all parts of the working project archive, it is unlikely that every digital file generated during a project will need to be retained for long-term preservation in the preserved archaeological archive. Any digital files created through the course of a project should therefore be subject to a selection strategy and the ClfA [Selection Toolkit for Archaeological Archives](#) provides practical advice to aid in the creation of sustainable archives (ClfA 2019).

How the digital archive is created, managed, and selected should not be a stand-alone operation, but should sit alongside other decision-making processes in the creation of the final archaeological archive. Reasons for selection of digital components will vary from project to project and should take into consideration the project aims and significance of the results, the research potential and contribution to research frameworks, and ability to provide greater access to heritage information.

Selection of archive material should be based on a process of consultation between the project team and relevant stakeholders – the ClfA Selection Toolkit offers general guidance on this.

This case study provides a fictional example of the types of digital material that can be created through the course of an archaeological project, and the selection processes that can be applied to the working project archive.

The Dig Digital resource is an Archaeological Archives Forum guidance document that supports ClfA Standards and guidance. It was created by DigVentures, in partnership with ClfA, and funded by Historic England.

You can find the full resource online at: <https://www.archaeologists.net/digdigital>



Approaches to selection

It is important to note that each project is unique, and the data management within the project will be determined by a series of factors specific to both that project and the circumstances of its delivery. The selection process for well-preserved site under full excavation will differ to that of a watching brief on a sterile site.

The way that we acquire, modify, manage and then select digital elements for inclusion in an archive will be affected by a range of factors that can change from project to project. Processes which determine how digital data is handled throughout the life of the project will be linked to these factors.

For example, a large-scale excavation project will produce huge quantities of both born-digital data and digitised records and it is likely that the resultant preserved digital archive will be sizeable, even where following a stringent selection process. By comparison, some archaeological projects, such as a watching brief or small-scale evaluation, are unlikely to produce a large digital data archive for deposition. On a site where no archaeological remains were recorded, a technical archaeological report attached to the OASIS record with a selection of images may suffice, if agreed as an appropriate means of archiving by all stakeholders as part of the selection strategy process.

Example

The table below provides an example of digital material relating to a fictional archaeological evaluation at St Cuthbert's Row in South Dreamshire. The [selection strategy](#) documenting this process for the complete archaeological archive can be found under the strategy examples tab of the [ClfA Selection Toolkit](#).

This fictional case study has been included to provide a practical and visual example of how the digital files created through the course of a project may be selected for long term preservation as part of a deposited digital archive. Files selected for inclusion in the archaeological archive are highlighted in **blue** and those combined from individual files into a single PDF/A format are highlighted in **yellow and underlined**.

Throughout project delivery, particular stages will provide review points where updates to the selection strategy and DMP can be made. These will ideally be linked to key stages that provide an opportunity for review, specialist input and evaluation, such as project planning, data gathering, analysis and reporting, and archive compilation.

This example describes some of the selection decisions made at each of the project review stages and the files selected for inclusion within in the preserved digital archaeological archive.



Archaeological evaluation at St Cuthbert's Row, Gt Lipington, South Dreamshire.		Project SD2367
/ Background Data / H&S	SD2367WSIfigure2V2-0.pdf / References Regional-Research-Framework-Imagination-Counties-2014.pdf SD2287GeophysicsReport2.0.pdf Medieval-Settlement-Gt.-Liping-Journal Article-2011.pdf SD2367SigninginSheet1-0.docx SD2367RiskEvaluationDraftV1-1.docx SD2367RiskEvaluationV2.0.docx SD2367Heath and Safety manualV3-2.pdf	Background documents and earlier site reports do not need to be included in the digital archive if they are properly referenced and available elsewhere.
<p>Data gathering</p> <p>The outline selection strategy should be monitored, reviewed and if necessary, amended through the course of the project in agreement with the relevant stakeholders. All members of the project team should be aware of the selection strategy and how it should be applied. In practical terms for the creation of digital material, this could mean training staff on data creation, folder hierarchies and in-house file-naming conventions, all of which should be detailed in the DMP. Therefore, the DMP and selection strategy should be available to all members of the project team, reviewed and amended as necessary.</p> <p>The archaeological evaluation at St Cuthbert's Row has created both born-digital data (data that is produced in digital form) and data that has been digitised (including records converted from paper or other analogue forms). While draft versions of digital files such as documents and illustrations were useful as part of the working project archive, only the final versions have been selected for inclusion in the preserved archaeological archive.</p> <p>Decisions around selection of digital material were based around accessibility and re-use: what data needs to be included so future researchers can re-examine and reinterpret the site. It is especially important that the data from this evaluation can be easily accessed and reused in future stages of the project. For this reason, the preserved digital archive includes the survey data and GIS shape files. The images, however, were subjected to a comprehensive selection process, removing all repeat, irrelevant or sub-standard files from the preserved archive.</p>		
/ Site Data / Registers	/Context Database SD2367AllContextDataWorkingCopyV1-1.xlsx SD2367AllContextDataWorkingCopyV1-1.xlsx /Matrices SD2367MatrixV1-0.xlsx SD2367MatrixV1-1.xlsx SD2367MatrixV2-0.xlsx SD2367MatrixV2-1FINAL.xlsx / Photographs SD2367PhotographsV1-0.xlsx SD2367PhotographsV1-1.xlsx SD2367PhotographsV2-0.xlsx / Drawings SD2367DrawingsV1-0.xlsx SD2367DrawingsV2-0.xlsx / Registered Finds	SD2367MatrixV2-1FINAL.xlsx Final version of Harris Matrix selected for archiving SD2367PhotographRegisterV2-0.csv SD2367DrawingRegisterV2-0.csv SD2367RegisteredFindsListV2-0.csv SD2367ContextRegisterV2-0.csv Final versions of all registers selected for archiving



Archaeological evaluation at St Cuthbert's Row, Gt Lipington, South Dreamshire.		Project SD2367
	<p>SD2367Registered-FindsV1-0.xlsx SD2367Registered-FindsV1-2.xlsx SD2367Registered-FindsV2-0.xlsx</p> <p>/ Contexts SD2367ContextV1-0.xlsx SD2367ContextV1-1.xlsx SD2367ContextV1-2.xlsx SD2367ContextV2-0.xlsx</p> <p>/ Photographs</p> <p>/ On Site JPG files: SD2367SitePhoto001.JPG to SD2367SitePhoto087.JPG</p> <p>CR2 files: SD2367SitePhoto001.CR2 to SD2367SitePhoto087.CR2</p> <p>/ Registered Finds JPG files: SD2367RS01V1-0.JPG SD2367RS01V1-1.JPG SD2367RS02V1-0.JPG SD2367RS02V1-1.JPG SD2367RS02V1-2.JPG SD2367RS03V1-0.JPG</p> <p>CR2 files: SD2367RS01V1-0.CR2 SD2367RS01V1-1.CR2 SD2367RS02V1-0.CR2 SD2367RS02V1-1.CR2 SD2367RS02V1-2.CR2</p>	<p>SD2367SitePhoto009.tiff to SD2367SitePhoto027.tiff SD2367SitePhoto030.tiff to SD2367SitePhoto045.tiff SD2367SitePhoto047.tiff to SD2367SitePhoto053.tiff SD2367SitePhoto065.tiff to SD2367SitePhoto085.tiff <i>Selected site photographs x 59 in preservation format (duplicate shots, blurred images and working shots de-selected)</i></p> <p>SD2367RS01V1-1.tiff SD2367RS02V1-2.tiff SD2367RS03V1-0.tiff <i>Selected registered finds photographs in preservation format</i></p>
	<p>/ GIS</p> <p>/ Survey SD2367GIS-TrenchOutlines.csv</p> <p>/ Rasters SD2367OSMap1906.pdf SD2367OSMap1906.pdf.aux.xml SD2287GeophysicsData.pdf SD2287GeophysicsData.pdf.aux.xml</p> <p>/ Shapefiles SD2367TrenchOutlines.cgp SD2367TrenchOutlines.dbf SD2367TrenchOutlines.prj SD2367TrenchOutlines.qpj SD2367TrenchOutlines.shp SD2367TrenchOutlines.shx</p>	<p>SD2367GIS-TrenchOutlines.csv</p> <p>SD2367TrenchOutlines.cgp SD2367TrenchOutlines.dbf SD2367TrenchOutlines.prj SD2367TrenchOutlines.qpj SD2367TrenchOutlines.shp SD2367TrenchOutlines.shx <i>All shapefiles selected for long-term preservation</i></p>

Analysis and reporting

The entire working project archive should be subject to a selection assessment in agreement with the relevant stakeholders. In terms of the digital data this should be informed by the DMP and deposition requirements of the TDR. During an archaeological project many versions of a document, spreadsheet or illustration can be created, and the evaluation at St Cuthbert's Row is no different. While draft versions of the report and finds management database were useful during analysis and reporting, they were considered to be of low intrinsic value, so only the final versions of these files have been retained for inclusion in the preserved archaeological archive.

To facilitate re-use, the data needs to be archived in an appropriate format. The project team have therefore included the 'raw materials' of what makes up the final report as well as a PDF/A version of the combined technical report (text, illustrations and tables) for easy dissemination. The preserved digital archive also includes the full specialist reports and original specialist analysis and data, as well as final versions of the illustrations and databases.

<p>/Reports</p>	<p>/ Eval Draft SD2367EvaluationReportV1-0.docx SD2367EvaluationReportV1-1.docx SD2367EvaluationReportV2-0.docx SD2367EvaluationReportV2-1.docx</p> <p>SD2367EvaluationReportV2-0.pdf <i>Combined PDF for circulation</i> SD2367EvalReportSDCCCommentsV2-0.pdf</p> <p>/ Eval Final SD2367EvalReportV2-2FINAL.docx</p> <p>Updated final PDF: SD2367EvalReportV2-2FINAL.pdf</p> <p>/ Evaluation Appendices SD2367Appendix1AllContextsTablesV1-0.docx SD2367Appendix1AllContextsTableV2-0.docx SD2367AllFindsReportsV1-0.docx SD2367AllFindsReportsV2-0.docx</p> <p>SD2367Appendix1AllContextsTableV2-0.pdf SD2367Appendix2AllFindsReportsV2-0.pdf</p>	<p><i>Generally, draft versions do not need to be included in the digital archive unless they contain specific/important revisions or comments.</i></p> <p>SD2367EvalReportV2-2FINAL.docx</p> <p>SD2367Appendix1AllContextsTableV2-0.docx SD2367Appendix2AllFindsReportsV2-0.docx</p> <p>SD2367EvalReportV2-2.pdf/a <i>Combined pdf text, illustrations and appendices saved as PDF/A for long-term preservation</i></p>
<p>/ Photographs</p>	<p>/ Eval Report Images JPG files: SD2367SitePhoto011Plate1.JPG SD2367SitePhoto033Plate5.JPG SD2367SitePhoto036Plate7.JPG SD2367SitePhoto042Plate2.JPG SD2367SitePhoto049Plate8.JPG SD2367SitePhoto053Plate4.JPG SD2367SitePhoto058Plate3.JPG SD2367SitePhoto067Plate9.JPG SD2367SitePhoto075Plate6.JPG</p> <p>CR2 files: SD2367SitePhoto011Plate1.CR2 SD2367SitePhoto033Plate5. CR2 SD2367SitePhoto036Plate7.CR2 SD2367SitePhoto042Plate2.CR2</p>	<p>SD2367SitePhoto011Plate1.tiff SD2367SitePhoto033Plate5.tiff SD2367SitePhoto036Plate7.tiff SD2367SitePhoto042Plate2.tiff SD2367SitePhoto049Plate8.tiff SD2367SitePhoto053Plate4.tiff SD2367SitePhoto058Plate3.tiff SD2367SitePhoto067Plate9.tiff SD2367SitePhoto075Plate6.tiff</p> <p><i>All raster images chosen for the report selected for inclusion in the archive in preservation format</i></p>

<p>/ Materials</p>	<p>SD2367SitePhoto049Plate8.CR2 SD2367SitePhoto053Plate4.CR2 SD2367SitePhoto058Plate3.CR2 SD2367SitePhoto067Plate9.CR2 SD2367SitePhoto075Plate6.CR2</p> <p>/ Finds data SD2367FindsManagementV1-0.xlsx SD2367FindsManagementV1-1.xlsx SD2367FindsManagementV1-2.xlsx</p> <p>/Specialists SD2367HarryKendtick-contract-letter.docx SD2367EmmaTurner-contract-letter.docx SD2367TamaraGreen-contract-letter.docx SD2367JemmaPalmer-contract-letter.docx SD2367DrFrancisBenjamin-contract-letter.docx</p> <p>/ Specialist Reports SD2367MedievalPotteryV1-1.docx SD2367MedievalPotteryV1-2.docx SD2367MedievalPotteryDataV1-0.xls SD2367MedievalPotteryMetadataV1-0.xls</p> <p>SD2376CBMV1-0.docx SD2376CBMtablesV1-0.xls SD2376CBMMetsdataV1-0.xls</p> <p>SD2367AnimalBoneReportV1-0.docx SD2367AnimalBoneReportV1-1.docx SD2367AnimalBoneDataV1-0.xls SD2367AnimalBoneMetadata.xls</p> <p>SD2367MetalWorkingV1-0.docx SD2367MetalworkingAnalysisV1-1.docx SD2367MWMetadata.docx</p> <p>SD2376PollenReportV1-0.docx SD2376PollendataV1-0.xls SD2367PollendataV1-1.xls SD2367PollenMetadataV1-0.csv</p>	<p>SD2367MedievalPotteryV1-2.docx SD2367MedievalPotteryDataV1-0.xls SD2367MedievalPotteryMetadataV1-0.csv</p> <p>SD2376CBMV1-0.docx SD2376CBMtablesV1-0.xls SD2376CBMMetsdataV1-0.csv</p> <p>SD2367AnimalBoneReportV1-1.docx SD2367AnimalBoneDataV1-0.csv SD2367AnimalBoneMetadataV1-0.csv</p> <p>SD2367MetalWorkingReportV1-1.docx SD2367MWMetadataV1-1.csv</p> <p>SD2376PollenReportV1-0.docx SD2367PollendataV1-1.xls SD2367PollenMetadataV1-0.csv</p> <p><i>Select all final versions of specialist reports, data and metadata</i></p>
<p>/ Illustrations</p>	<p>/ Eval Drafts SD2367Figure1V1-0.ai SD2367Figure1V1-1.ai SD2367Figure2V1-0.ai SD2367Figure3V1-0.ai SD2367Figure3V1-1.ai SD2367Figure3V1-2.ai SD2367Figure4V1-0.ai</p> <p>/ Eval Final SD2367Figure1V2-0.ai SD2367Figure2V2-0.ai SD2367Figure3V2-0.ai SD2367Figure4V2-0.ai</p> <p>SD2367Figure1V2-0.pdf SD2367Figure2V2-0.pdf</p>	<p>SD2367Figure1V2-0.svg SD2367Figure2V2-0.svg SD2367Figure3V2-0.svg SD2367Figure4V2-0.svg</p> <p><i>All vector images chosen for the report selected for inclusion in the archive in preservation format</i></p> <p>SD2367FiguresCombinedV2-0.pdf/a</p>

	SD2367Figure3V2-0.pdf SD2367Figure4V2-0.pdf	<p>Combined pdf of vector illustrations saved as PDF/A for long-term preservation</p>
<p>Archive compilation</p> <p>The final version of the selection strategy (including all amendments and revisions), relevant repository contracts and project metadata have been included in the digital archive to be deposited with the trusted digital repository for long-term preservation. In a similar way to how the context or drawing registers function for other parts of the archive, the metadata tables provide a register and field-level definition of the data collected. While they were finalised during archive compilation, the metadata tables were started earlier in the project and amended as necessary.</p> <p>General project admin and internal documents have not been selected for inclusion in the preserved digital archive, though they will be retained on the excavating unit's servers for future reference.</p>		
<p>/ Admin</p>	<p>/ Tender T2016-023GuthbertsRowTenderV1-0.docx T2016-023rGuthbertsRowTendeV1-0.pdf</p> <p>/ Finances SD2367BudgetV1-0.xls SD2367BudgetV1-1.xls SD2367PaymentScheduleV1-0.xls</p> <p>/ Delivery SD2367ClientContractV1-0.docx SD2367ClientContractV1-1signed.docx SD2367GanntV1-0.xls SD2367GanntV1-1.xls</p>	
<p>/ Archive</p>	<p>/ Museum Agreements SD2367SDMSCContractV1-0.docx SD2367SDMCCContractV2-0.pdf SD2367SDMSTransferTitleV1-0.pdf SD2367SDMSTransferTitleSignedV2-0.pdf SD2367SelectionStrategyFINAL.docx</p> <p>/ Data Management Plan SD2367DMPV1-0.docx SD2367DMPV1.1.docx SD2367DMPV1-0.docx SD2367DMPV2-1FINAL.docx</p> <p>/ Project Metadata SD2367CollectionLevelMetadataV1-0.doc SD2367CollectionLevelMetadataV1-2.doc</p> <p>SD2367FileLevelMetadataV1-0.xls</p>	<p>SD2367SDMCCContractV2-0.pdf SD2367SDMSTransferTitleSignedV2-0.pdf SD2367SelectionStrategyFINAL.docx</p> <p><i>The final version of the selection strategy should be included in the preserved archive in addition to the outline strategy agreed during project planning</i></p> <p>SD2367DMPV2-1FINAL.docx</p> <p>SD2367CollectionLevelMetadataV1-2.doc</p> <p><i>File-level metadata were developed as an .xls spreadsheet through delivery, and individual .txt files will be created from this for submission of data to accompany each element deposited.</i></p>