Research Data/Digital outputs underpinning theses: deposit in Pure

March 2018

Research Data/Digital outputs underpinning theses

Contents

Data categories

Below are some helpful definitions that can help define the scope of the requirements.

Primary underpinning data/digital outputs: data or digital outputs that are the foundation of your thesis but are not essential to the reader of your theses to understand it. These type of data or digital outputs should be submitted by following the guidelines in the pages below.

Primary supplementary or complementary data/digital outputs: data or digital outputs that are an integral part of your thesis, they are essential to the reader to understand your thesis. These type of data or digital outputs should be submitted to registry alongside your thesis.

Secondary data/digital outputs: data or digital outputs that were collected or generated by someone other than you. Examples are: census data, government data, electoral statistics and NHS data.

Personal/sensitive/confidential data or digital outputs: if your data can be describe by this category, please get in touch with us (research-data@st-andrews.ac.uk) as you might need to request an embargo.

If you are unsure whether your data fall under one of the above categories, please email us at research-data@st-andrews.ac.uk and we will be happy to help.
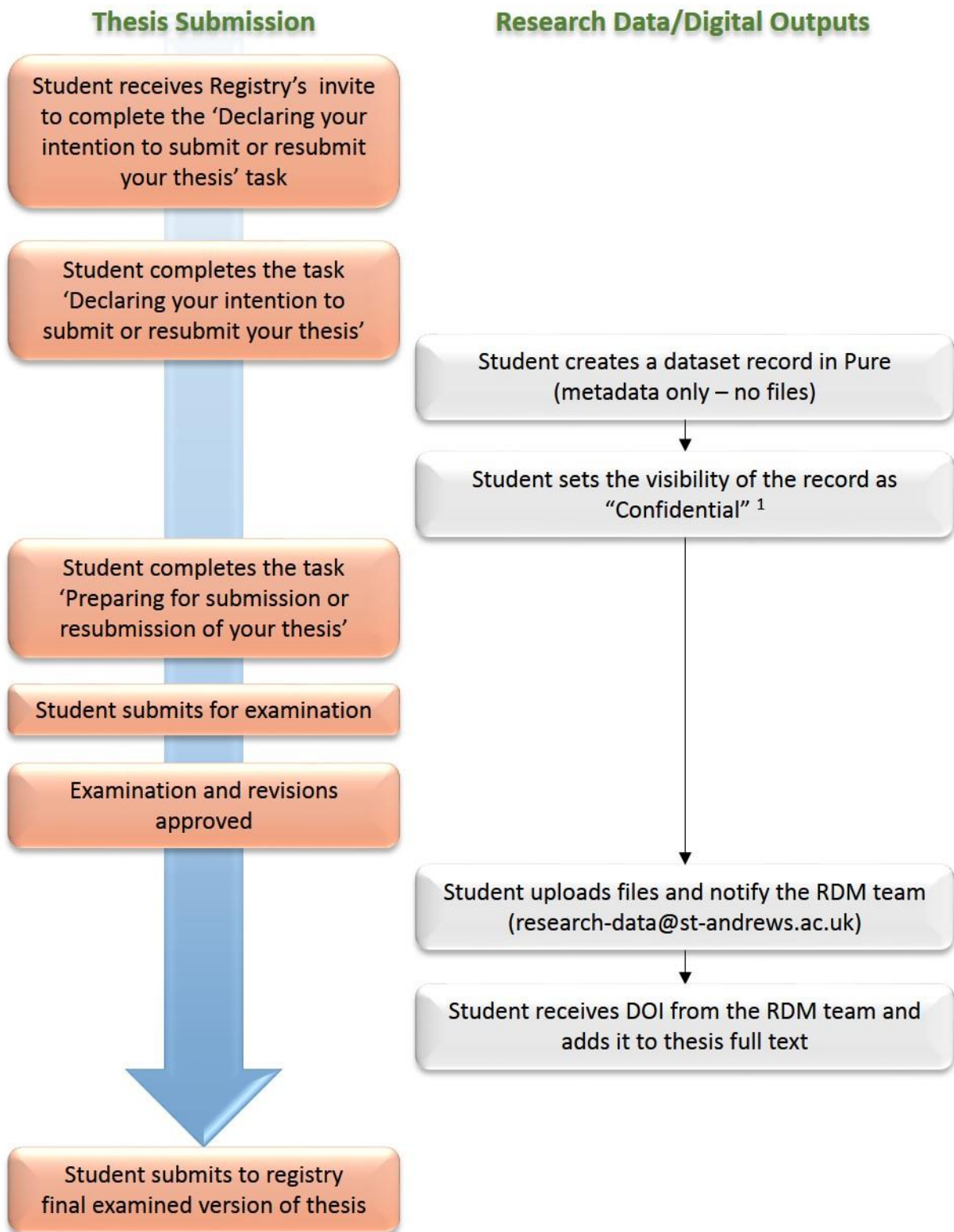
When to deposit

Below is the sequence of steps that we advise you to follow in order to deposit your dataset in Pure and obtain a DOI. See p. 2 for the workflow referenced against the thesis submission process. Please note that the thesis submission process illustrated in this document is not to be used as guidance for the submission of your thesis.

Before submitting your thesis for examination:

1. Create a basic dataset record in Pure:

    o   Provide at least Title, Authors and Date made available (the year will be enough, The Library will fill in the rest of the date) (p. 3-5)

    o   Set the metadata visibility as Backend – since the record is in preparation it should not be publicly visible yet (p. 5)

Before submitting to registry the final examined version of your thesis:

2. Upload your final data files in Pure (for more details on file visibility see p. 4) and email research-data@st-andrews.ac.uk. **IMPORTANT:** if you have requested an embargo for your data the files will not be released until both supervisor and student give their permission.

3. We will send you (and your supervisor) a DOI for your dataset.

4. Add the DOI to your thesis full text using the acknowledgement template available in MySaint (same location as your tasks). If you need advice, please, email us at research-data@st-andrews.ac.uk.

Thesis Submission

Research Data/Digital Outputs

Student receives Registry's invite to complete the 'Declaring your intention to submit or resubmit your thesis' task

Student completes the task 'Declaring your intention to submit or resubmit your thesis'

Student creates a dataset record in Pure (metadata only – no files)

Student sets the visibility of the record as "Confidential" [1]

Student completes the task 'Preparing for submission or resubmission of your thesis'

Student submits for examination

Examination and revisions approved

Student uploads files and notify the RDM team (research-data@st-andrews.ac.uk)

Student receives DOI from the RDM team and adds it to thesis full text

Student submits to registry final examined version of thesis

[1]User Guide: "Depositing datasets in Pure: when and how", p.5

University of St Andrews | FOUNDED 1413

Email: research-data@st-andrews.ac.uk
Phone: 01334 462343

University Library

How to deposit

Log in into Pure at https://www.st-andrews.ac.uk/staff/research/pure/. From your personal profile you can open a Datasets template by clicking on **Add New** (on the right-hand side of the screen) and selecting **Datasets** from the **Choose submission** window; or by clicking on the '**+**' beside 'Datasets' on the left-hand side menu

OR
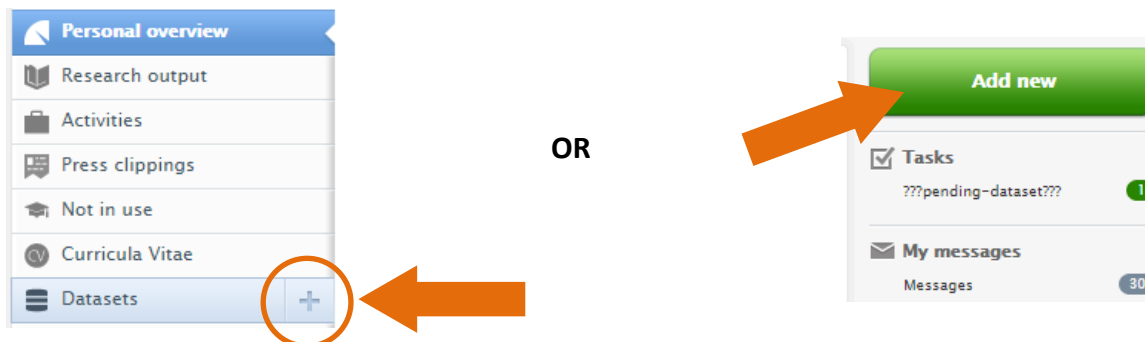
Identification

Title: We advise to use the form: "*Title* **(thesis data)***"*.

Description: you should briefly describe the dataset, what type of files are attached and in which formats, if proprietary software is required (provide details) and any other relevant information that can help the user understand the datasets. More detailed information could also be provided as a ReadMe.txt file alongside the data files. See appendix for more details and examples.

Temporal coverage: it indicates the date range coverage of the data, for example data covers housing records from 1850-1905. This field is not mandatory.

Date of data production: this can either be a specific date or a period of time over which the data have been collected/produced. Please note that this field becomes mandatory only if one of the two options is selected. This field is not mandatory.

People

This will be pre-populated with your details. You can add your supervisor but please remember to change their role to 'Supervisor' in the drop down list for 'Role'.

Search for authors that are internal to the University

Create an author from a different institution or organisation

Search and add Person – or create External Person

Search   Create external person

First name *

Last name *

Role *
Creator

Type
External person

Country

Managing organisational unit

This sections will be pre-populated with your details.

Data availability

Publisher: this is set by default to "University of St Andrews". This field does not require any action from you.

DOI: the Library will email you a DOI when you upload the files. The Library will also add the DOI to the Pure record.

Electronic data: this is where the files can be uploaded. **Note: If your data are of sensitive and/or confidential nature, emails us (**research-data@st-andrews.ac.uk**) before uploading the files. If unsure, we can help clarifying.**

Data availability

Publisher *
🏛 **University of St Andrews**
Publisher

Change publisher...

DOI

Add existing DOI   –or–   If you do not have a DOI already one will be issued during validation where appropriate

Electronic data

☁ Drag files or **browse** your computer.
Max filesize: 1024 MB.

Physical data
Add details of physical data...

Links
Add link...

Date made available
Year *   Month   Day

 After having selected the file(s), the 'Multiple file upload' window will open. Note that max file size is currently set at 2GB. For larger files, please, contact us. You should only upload individual files or compressed folders. Your files can be organised in compressed folders by chapter, analytical technique or any other criteria that you deem suitable.

		Multiple file upload			
File name	File size	Visibility	License	Type	Embargoed until
Test_file_2.docx	11 KB	Public – No restriction	CC BY	Audiovisual	
Test_file_1.docx	11 KB	Public – No restriction	CC BY	Audiovisual	

Visibility – you can set this as 'Backend'. The Library will always check that the correct setting are applied.

License – Please, choose a licence as this will inform the user on the terms and conditions of use. You can either use Creative Commons or Open Data Commons licences. For more information contact us or visit http://creativecommons.org/ and http://opendatacommons.org/.

Type – make sure that the appropriate type is selected. If you are unsure, please select '*Dataset'.*

Embargoed until – the Library will set the correct embargo based on the embargo requested in the thesis declaration.

Physical data: here you can describe any physical item that can be made available upon request (e.g.: samples, tissues, notebooks, etc.). This field is not mandatory.

Date made available: you can simply add the year and the Library will fill in the rest when the files will become publicly available.

Access option - For Library use only

Relation to content

In this section you will be able to make links to other items in Pure such as articles that you have published, activities, posters and others. You can create a link by clicking on the "+" icon and searching by Title or Pure ID. If you need any assistance, please get in touch.

Visibility (metadata)

This will dictate if the metadata (general information: Title, description, authors,

Visibility ⓘ

○ Public – No restriction

○ Public – No restriction
● Campus – Restricted to specific IP range
● Backend – Restricted to Pure users
● Confidential – Restricted to associated users and editors

publisher, etc.) will be publicly visible on the University's portal. By default this is set to Public.
IMPORTANT: You should set this to 'Confidential'.

Status

By default this is set as "For validation", you should change it to "Entry in progress".

○ Pu

PURE REGISTRATION STATUS
Entry in Progress
✓ For validation

Status: For validation ▼

Save

Remember to save when done or all the changes will be lost

Save

University Advice and Guidance

University RDM website	https://www.st-andrews.ac.uk/library/services/researchsupport/researchdata/
Contact us	Email - research-data@st-andrews.ac.uk ; Phone – 01334 46**2343**
Licences	http://creativecommons.org/ ; http://opendatacommons.org/
Pure	https://www.st-andrews.ac.uk/staff/research/pure/
Research Portal	https://risweb.st-andrews.ac.uk/portal/
General Pure enquiries	Pure Live System - purelive@st-andrews.ac.uk

Appendix A: example of description and ReadMe.txt file

With your data you should provide enough information to give some context to the user. This can take the form of a simple description (in the "Description" field), a more comprehensive ReadMe.txt file or both. You should include basic information on the methodology employed to acquire the data (online surveys, paper surveys, instrumental analyses, etc.) together with type of files, their formats and what software is required to open, visualise and analyse them. If you have any questions or would like some help, please do not hesitate to get in touch.

Example of Description:

The following file types and formats are included:
- X-ray photoelectron spectroscopy files: .vms (can be opened with a text editor)
- X-ray diffraction files: .xrdml (proprietary format) and .udf (can be opened with a text editor)
- UV-Vis spectroscopy files: .jws (proprietary format) and .txt (can be opened with a text editor)
- Transmission electron microscopy: .dm3 (proprietary format) and .jpg
More details on the software required are provided in the ReadMe.txt file.

Example of ReadMe.txt file

The X-ray Photoelectron Spectroscopy (XPS) data have been acquired using a Kratos AXIS Ultra DLD, with a monochromatic Al Ka (1486.6 eV) X-ray source. The files are in .vms format and can be opened with any text file handling software (Notepad, Word etc.). The proprietary software that would allow for the visualisation of the data and their analysis is CasaXPS (http://www.casaxps.com/).

The powder X-ray diffraction data were collected using an Empyrean PANalytical series 2 diffractometer with a Cu Ka radiation source (l = 1.5406 Å). The files are offered in two different formats:
- .xrdml, an xml coding that can be opened with the Notepad application and with the PANalytical software;
- .udf, that can be opened with any text handling software. It contains details of the analysis and the intensity counts.
To perform analysis on this data files software such as PANalytical may be required.

The UV-Vis spectra were acquired using JASCO-V-650 spectrophotometer. The files are in both proprietary file (.jws) and text file (.txt). The proprietary format requires the software SpectraManager by JASCO to be opened this software will also allow to perform further analysis of the data.

Transmission electron microscopy images were acquired using a Jeol JEM-2011 operated at 200 KV using a Gatan 794 CCD camera to record digital images. The original images are in .dm3 format which can be visualised with the free software ImageJ but require the Gatan software to be analysed. Images in .jpeg format have also been included.

Appendix B: example of Research Data/Digital Outputs access statement ▬

Below are examples of how to include an access statement in your thesis. You can find a blank template in the in MySaint (same location as your tasks).

Research data underpinning this thesis are available at https://doi.org/10.4225/15/55D51A334B70777

OR

Digital outputs underpinning this thesis are available at https://doi.org/10.4225/15/55D51A334B70777