File Formats

There is no perfect file format. Each will have advantages and disadvantages depending on your research uses. Select a file format, or set of file formats, that helps you complete your research now, and that you can access again in the future. This is important both for your research outputs (what you create) and your research inputs (materials you use in the research process).

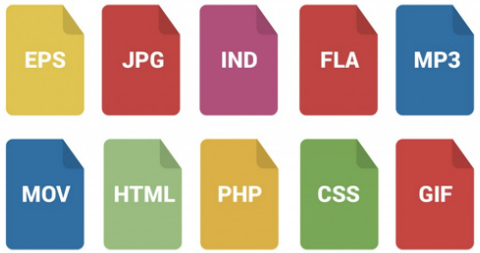
Common file types include:

-Images: jpg, gif, tiff, png, ai, svg, ...
-Video: mpeg, m2tvs, flv, dv, ...
-GIS: kml, dxf, shp, tiff, ...
-CAD: dxf, dwg, pdf, ...
-Data: csv, mdf, fp, spv, xlx, tsv, ...
-Text: txt, rtf, tvi, doc, pdf...

Consider what might happen if you can no longer use your software. Whether the software publisher goes bankrupt, the latest version refuses to read older data, or you can't afford a personal license for it after you graduate, the end result is the same. Losing access to your software can mean losing your data, especially if it is the only software that can read your data.

How to select file formats:

-Use software that imports and exports data in common formats.
-Ask advisors and colleagues what formats they use.
-Choose a format with functions that support your research needs.
-Save final versions of your content in multiple formats in order to spread your risk across multiple software platforms (e.g., docx, pdf, and txt; or mp4, avi, and mpg).



If you use website-based materials as evidence or references, take precautions to ensure that if the content moves, changes, or disappears, you still have evidence of its existence. Current tools to help you ensure the longevity of these materials include Robust Links and Archive-It. You can also take screenshots of important digital content in order to preserve the look and feel of an object.

Many ETD programs favor pdf files. If you export your research outputs to pdf, make sure you:

1. Embed your fonts
2. Embed (and test!) hyperlinks
3. Stabilize your web-based resources and citations (using a tool like Robust Links, Archive-It, or PermaCC)
4. Store supplementary materials as separate files
5. Verify the PDF/A compliance (use Acrobat Pro "Preflight" feature under "Edit")

Before you undertake any conversion, you need to identify what characteristics of your data are important to maintain during the conversion. For example, are the colors in a document or image important? Is the pagination essential? What about references? You will want to test these after your conversion is complete to ensure that you have a conversion that will meet your needs.

Additional Resources:

-List of File Formats (Wikipedia)
-Recommended Formats Statement (Library of Congress)
-Evaluating Your File Formats (UK National Archives)
-Reformatting Guides (US National Archives)

INSTITUTE of Museum and Library SERVICES

Source - Guidance Briefs: Managing Your ETD Research Files

EDUCOPIA INSTITUTE