A Brief Guide: Data Curation

CREATING
Good curation begins with developing strategies for collecting data, as well as creating documentation and metadata to explain and contextualize them. Data curators can help create and review strategies.

MAINTAINING
Ongoing data maintenance supports quality assurance. Practices include implementing a secure storage plan and protective measures such as file backups and file integrity checks.

REUSING
To encourage reuse, data curators may augment metadata, transform file formats to improve interoperability, and link to related datasets to increase discoverability.

USING
Applying robust quality control and assurance practices during data analysis -- such as file versioning and metadata review -- facilitates and expedites data archiving and reuse.

SHARING
Data curators facilitate the deposit of data, code, and metadata into reputable repositories that enable discovery and access. This includes applying appropriate anonymization or access controls.

ARCHIVING
Data curators select data for long-term preservation and then facilitate bundling data, code, and metadata into a single package for transfer to an archive.

Data curation is the active management of research data throughout a project to produce datasets that are FAIR: Findable, Accessible, Interoperable, and Reusable.

This guide was created by the Portage Curation Expert Group and can be modified and re-used freely under the CC-BY license.