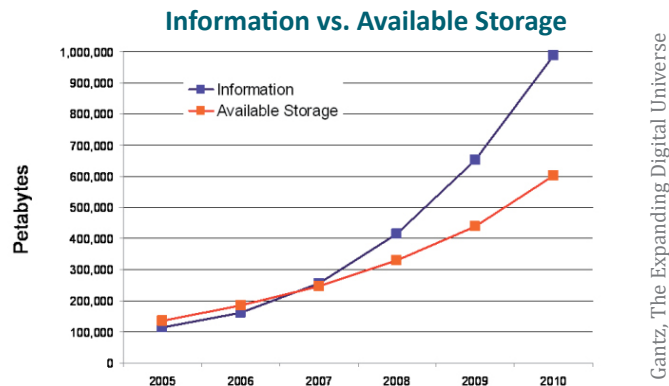


The world of data around us

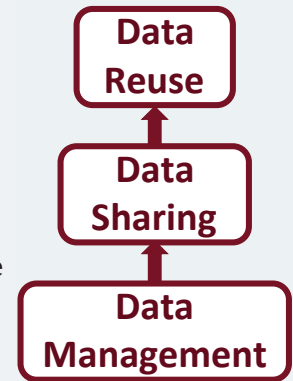
The data deluge has created a surge of information that needs to be well-managed, discoverable, and accessible.

The amount of available storage is not keeping pace with the amount of data being produced.



Why manage data: the researcher perspective

- Keep yourself organized ⇒ find your own files!
- Track your processes for reproducibility
- Better version control of data
- More efficient data quality control
- More backups to avoid data loss
- Format your data for reuse by yourself & others
- Document your data for understandability and reuse
- Prepare it to share it & gain credibility and recognition for your scientific efforts



Data management facilitates sharing and reuse.

Causes of data loss

- Natural disasters
- Facilities infrastructure failures
- Storage failure
- Server hardware or software failure
- Application software failure
- Human errors
- Malicious attack
- Format obsolescence
- Loss of competencies
- Loss of funding
- Loss of insitutional commitment

Costs of not doing data Storage management can be very high!

Data Reuse Example

Researchers reused and aggregated data from several different sources to determine migration routes for specific bird species.



© Jerry & Sherry Liguori

The Case for Data Management

If data are:

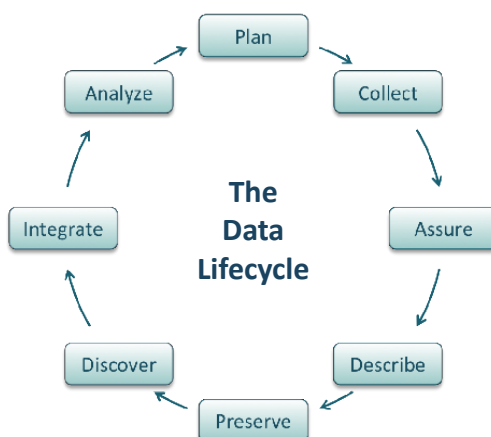
- Well-organized
- Documented
- Preserved
- Accessible
- Verified as to accuracy & validity

The results are:

- High quality data
- Data that is easy to share and reuse
- Citation & credibility to researcher
- Cost savings to further science

The Data Lifecycle

The stages through which well-managed data passes from project inception to conclusion.



Local contact information

