Lesson 6: Protecting Your Data



View all Education Modules at https://www.dataone.org/education-modules (CC)

Backups vs.

• Periodic snapshots of current version

- Stored for short or near-long-term
- Often done on a somewhat frequent schedule
- Final version for
 - historical reference or disasters

Archiving

- Stored for long-term Created at end of •
- project or at major milestone

Keep multiple copies of your backups & archives in case one copy fails!

Backups: Questions to consider

- Are there existing policies that might affect how • and when you do backups?
- How often should you do backups?
- Should you do partial or full backups? •
- What will you do with non-digital content?
- Where will you backup your files? •
- Will backups be manual or automatic? •
- How do you recover files from your backup?
- How will you verify that a backup has been per-• formed successfully?
- Do you need to create backups of the backups?
- How long will you keep your backups? •
- What will happen to your backups when funding • ceases, the project ends, or staff leave?
- Can you read data off of older backups?
- How will outdated data be disposed of?

Only back up the data you can't afford to lose!

Data in real life

A design firm's backup system was working fine. The backup software reported that backups were successfully completed. The administrator checked the backups right after they were done and confirmed they were good.

After a virus erased most of their files, they went to their backups and found that they were all blank. An investigation revealed that the backup tapes were stored near an elevator. Whenever the elevator went past, magnets inside it erased the tapes.

Had they checked their backups properly, they may have noticed this before there was an emergency.

Why perform backups?

- Limit or negate data loss, particularly of data that is not reproducible
- Save time and money
- Help prepare for disasters
- Reproduce results of past procedures
- Respond to data requrests
- Limit liability



image by Sybren A. Stüvel

Definition: Data preservation

Includes backups and archiving as well as processes such as data conversion, reformatting, and rescue.

Preservation: Questions to consider

- Are you using non-proprietary, standard formats?
- Are your file versions well-labeled?
- Are file names consistent, descriptive, & concise?
- Have you created a good data management plan?

Best Practices

- Create a detailed and thorough backup policy.
- Review your backup policy and plan periodically to ensure it is still valid and applicable.
- Minimize or remove reliance on users to perform manual backups.
- Don't assume backups are being made for you.
- Use non-proprietary, standard formats. •
- Check your backups manually. •
- Have multiple versions of backups on multiple • formats in multiple places.
- Practice good data management to limit the amount of data rescue that needs to be done to older data.

Local contact information