Paying for Long-Term Data Storage

Robin Dean

robin@coalliance.org

Digital Repository Director
Colorado Alliance of Research Libraries

How do researchers want to pay for data storage?

Bill them once

The grant pays for it

You keep their data . . .

FOREVER!

How does this work?

DataSpace @ Princeton

Pay Once, Store Forever
Write Once, Read Forever

$6 / GB
- Includes only the cost of storage
- Assumes storage cost decreases 20% every year

DataSpace Funding and Operational Model
DataSpace Policies and Guidelines

tDAR: the Digital Archaeological Record

"a modest one-time fee"

"permanent shared access"

"preserve files in perpetuity"

$2,500 / GB = 100 [10 MB] files

http://www.tdar.org/about/pricing/

Our reality . . .

We pay for infrastructure* every year

*primary storage and backup on networked hard drives that are always available

Also…

Costs do not decrease significantly on a year-to-year basis

David Rosenthal's blog on cloud storage costs

Therefore…

We charge for infrastructure every year

$25 / GB
- 1 primary copy (hosted hardware)
- 2 backup copies (hardware managed by us)
- 1 copy in DuraCloud (Amazon)

Who do we charge?

Libraries!

- Organizational home of digital repositories

- Have recurring funding (unlike researchers?)

Another model . . .

figshare

1 GB = free

10 GB = $80/yr

Recurring cost paid by researchers

Charge once . . .

- If you have additional funding from federal grants, foundations, or companies

- If each researcher's files are small enough that they can afford a realistic one-time charge for 100+ years of storage

- If you can count on a significant decrease in storage costs every year

Charge annually . . .

- If your business model requires you to cover the recurring costs for staff and facilities to manage storage and files

- If the researcher's files are large enough that a one-time cost would be prohibitive (> 1 GB?)

- If you store data on networked disks in a frequently accessed system (as opposed to tapes/offline storage)

- If you are paying someone else annually for infrastructure as a service!

Paying for
Long-Term Data Storage

Robin Dean
robin@coalliance.org

Digital Repository Director
Colorado Alliance of Research Libraries