

What Goes Where When in Managing Research Data?

David B. Lowe

Data Management Services Librarian
UConn Libraries

4/9/2014

University of Massachusetts and New England Area
e-Science Symposium

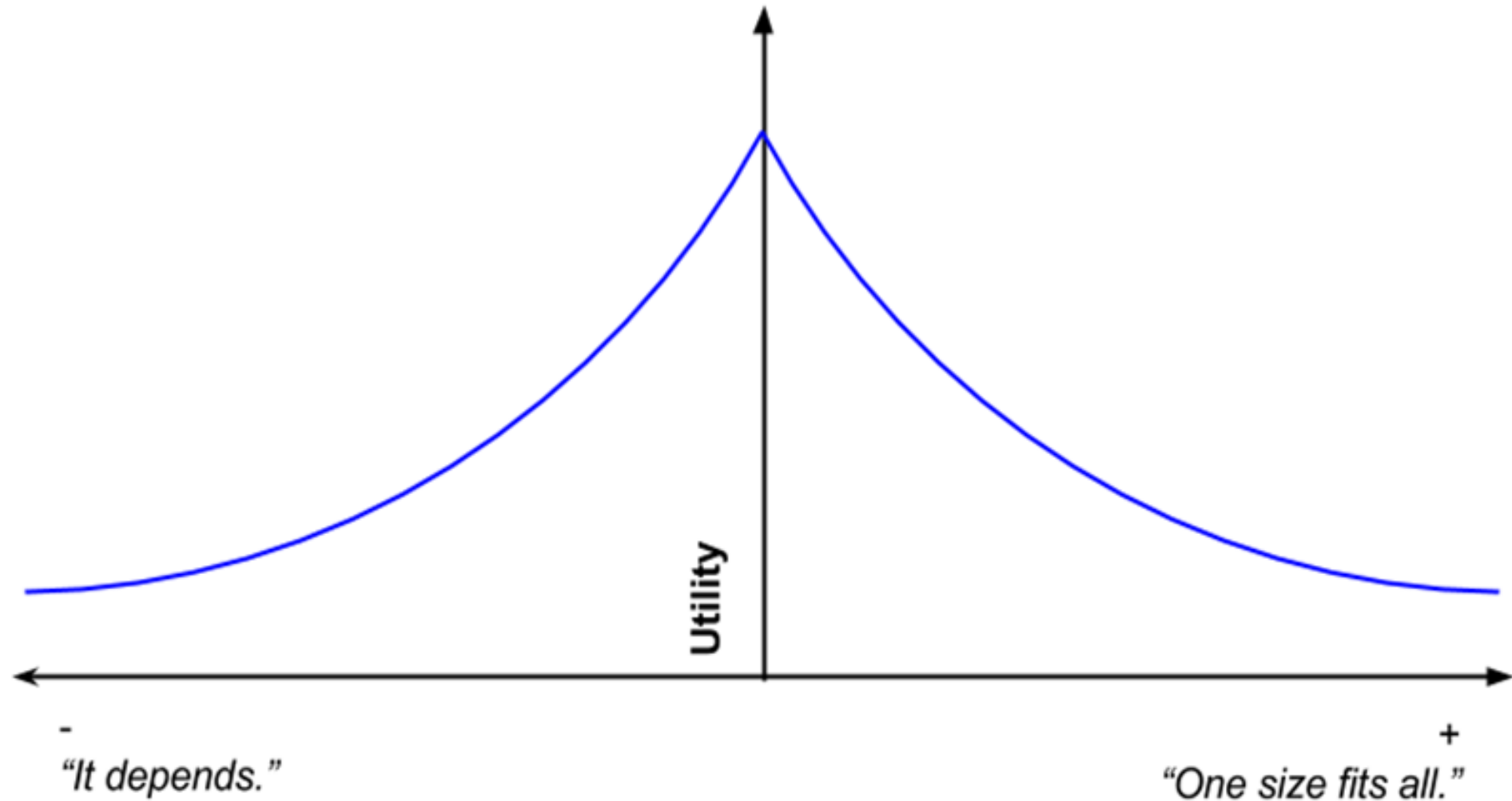
The Current State:

	NSF-AGS	NSF-AST	NSF-BIO	NSF-CHE	NSF-CISE	NSF-DMR	NSF-EAR	NSF-EFRI	NSF-EHR	NSF-ENG	NSF-GEN	NSF-PHY	NSF-SBE	Superset c
1	Products of research	Products of research	Data Collected, Formats, and Standards	Products of research	Roles and responsibilities	Types of data produced	Types of data	Types of data produced	Data generated by the project	Roles and responsibilities	Types of data produced	Types of data produced	Roles and responsibilities	Roles and responsibilities
2	Data format	Data format	Data Storage and Preservation	Data format	Types of data	Data and metadata standards	Data and metadata standards	Data and metadata formats	Period of data retention	Expected data	Data and metadata standards	Data and metadata standards	Expected data	Types of data produced
3	Access to data, and data sharing practices and policies	Access to data and data sharing practices and policies	Dissemination Methods	Access to data and data sharing practices and policies	Policies for access and sharing and appropriate protection/privacy	Policies for access and sharing, and provisions for appropriate protection/privacy	Policies for access and sharing	Access policies and provision	Data format and dissemination	Period of data retention	Policies for access and sharing	Policies for access, and sharing, and provisions for appropriate protection/privacy	Period of data retention	Period of data retention
4	Policies and provisions for re-use, re-distribution and production of derivatives	Policies and provisions for re-use, re-distribution and production of derivatives	Policies for Data Sharing and Public Access	Policies and provision for re-use, re-distribution and products of derivatives	Data storage and preservation of access	Policies and provisions for re-use, re-distribution	Policies and provisions for re-use, re-distribution	Means of sharing the outcome	Data storage and preservation of access	Data formats and metadata	Policies for re-use, redistribution	Policies and provisions for re-use, re-distribution	Data format and dissemination	Data format and dissemination
5	Archiving of data	Archiving of data	Roles and Responsibilities	Archiving of data	Additional possible data management requirements	Plans for archiving and preservation	Plans for archiving and preservation of access		Additional possible data management requirements	Data dissemination and policies for public access, sharing and publication delays	Plans for archiving and preservation	Plans for archiving and preservation	Data storage and preservation of access	Data storage and preservation of access
										Data storage and preservation of access			Additional possible data management requirements	Policies for access and sharing, and provisions for appropriate protection/privacy
														Policies and provisions for re-use, re-distribution
														7 Plans for archiving and preservation
														8 on
														9 Dissemination
														10 Additional possible data management requirements

So do you feel like this?

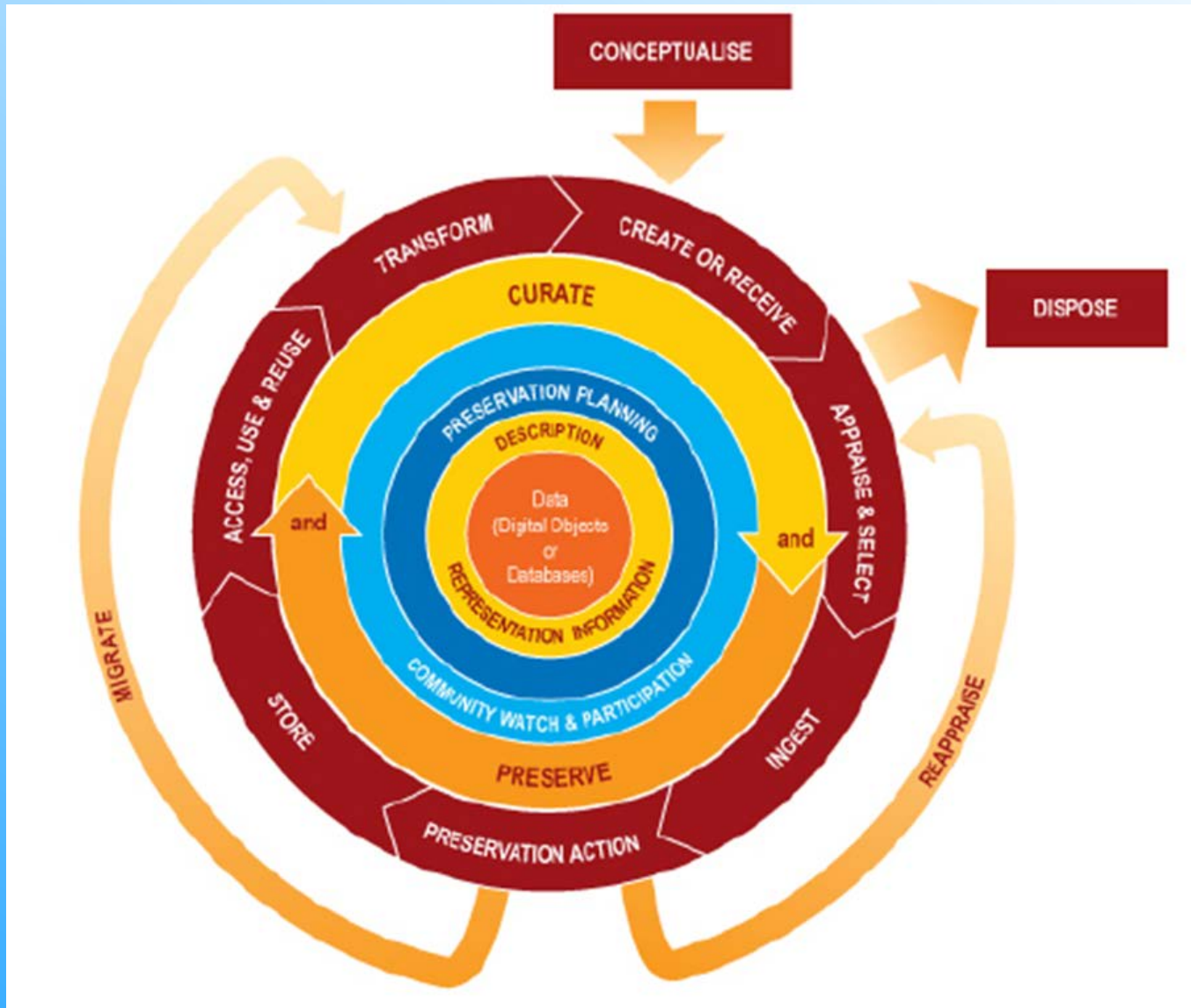


Utility of Suggested Solutions in the Context of Innovative Environments

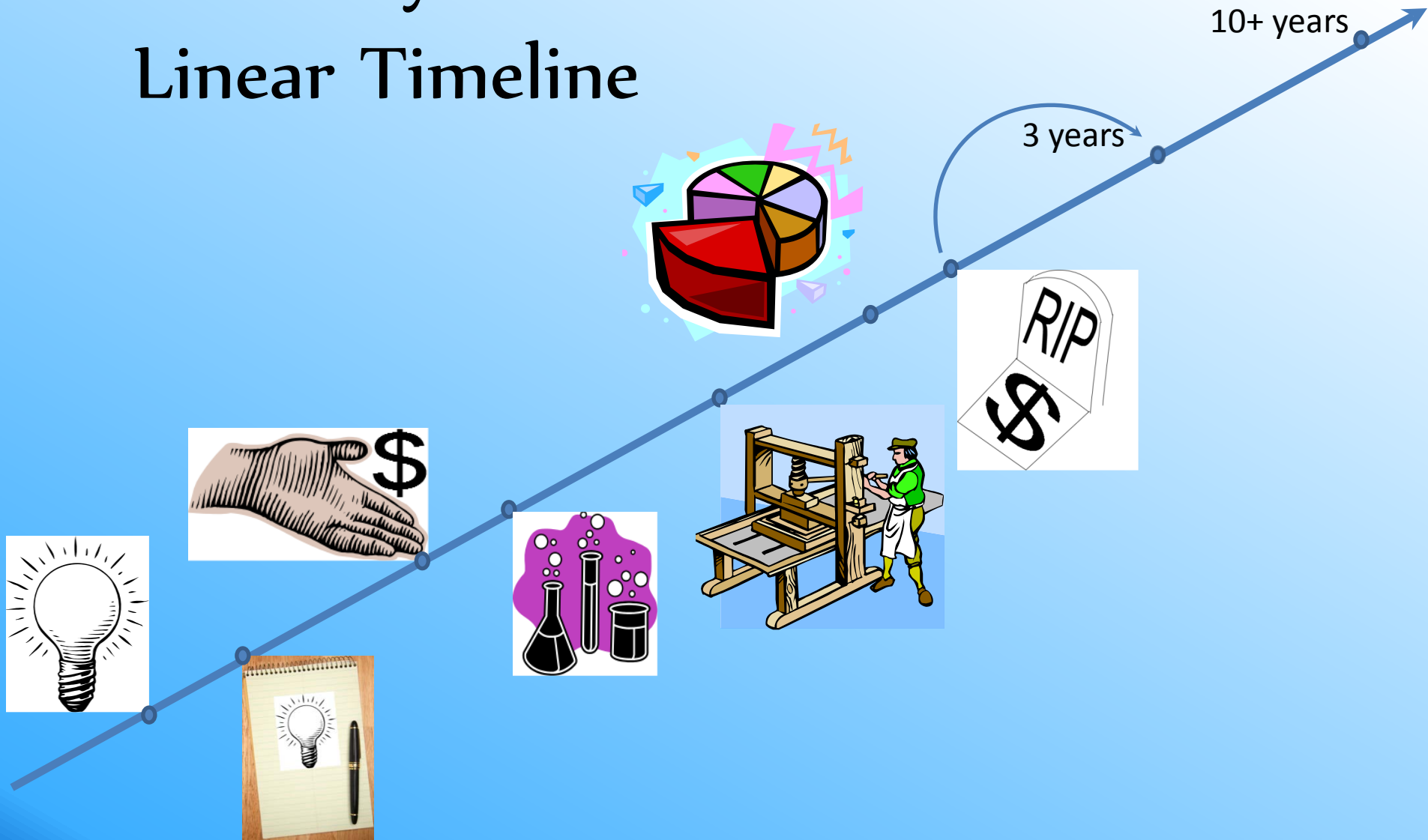


Data Life Cycle Model

(Digital Curation Centre, UK)



From Cycle to Linear Timeline



Data workspace storage options available to UConn affiliates as of 3/2014

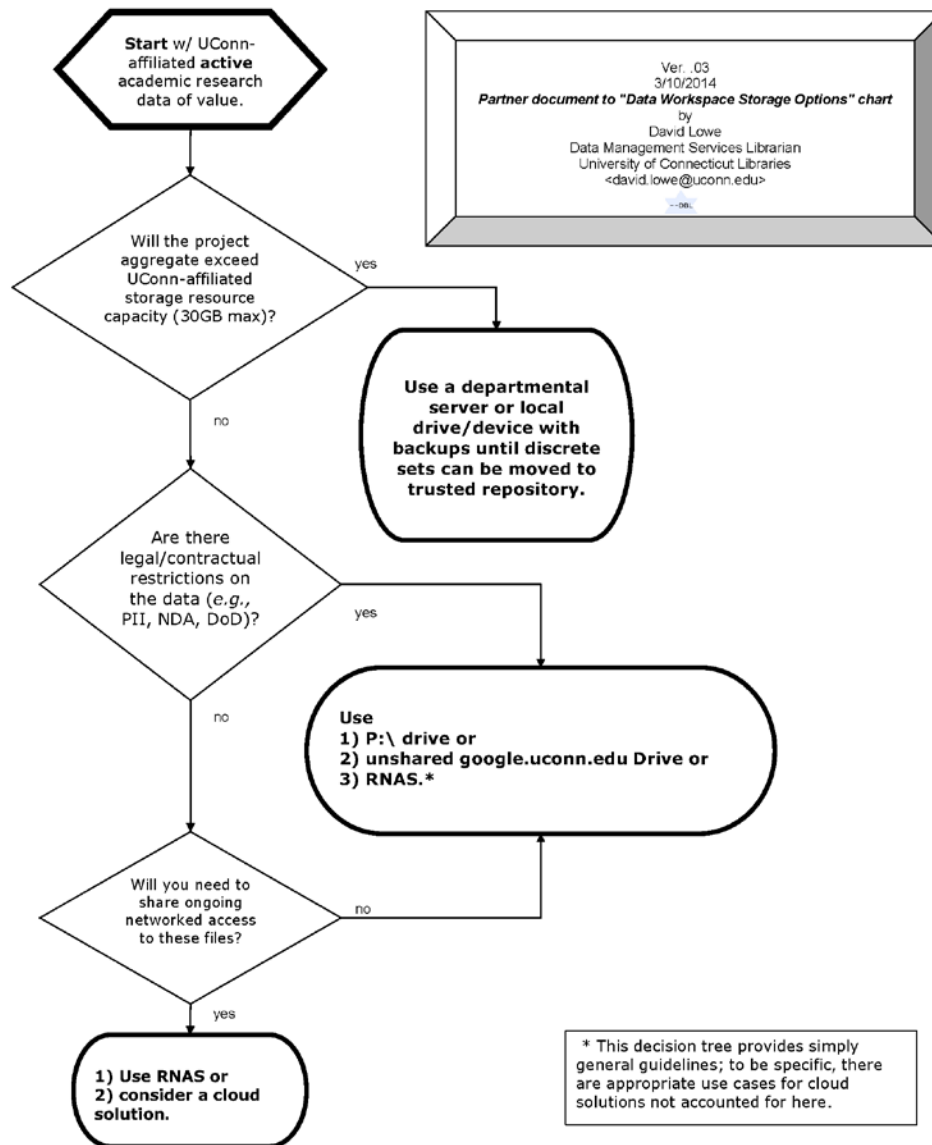
I. UConn Affiliate Resource:	maxGB free	GroupShareable (within UConn)	GlobalShareable (beyond UConn)	☺	☹	Link:
ECM/FileNet	TBD	Yes	No	secure	State-hosted; not single-sign-on; more for dept. documents	http://ecm.uconn.edu
EFS or "P:\ drive," individual	10	No	No	secure	questionable long term solution	http://fileserver.uconn.edu
EFS or "P:\ drive," departmental	10/member	Yes	No	secure	questionable long term solution	"
filelocker.uconn.edu	2	Yes	Yes	secure	short-term storage, sharing	http://filelocker.uconn.edu
github.uconn.edu	1	Yes	Yes	secure	not for large or A/V files, but code	http://github.uconn.edu
google.uconn.edu Docs/Drive	30	Yes	Yes	secure	not long-term solution	http://google.uconn.edu
RNAS storage device	15	Yes	Yes	secure	now B2B only via dept.; B2C coming; not free above allotment	http://rmas.uconn.edu/
II. Departmental or Personal Resource:						
Departmental server	varies	varies	varies	convenient	questionable long term solution	{your department}
media/devices	varies	No	No	convenient	highly fragile	{your favorite electronics retailer}
III. Cloud solutions:						
Box.com	10	Yes	Yes	convenient	terms/security questionable	http://www.box.com/
Dropbox.com	2+	Yes	Yes	convenient	terms/security questionable	http://www.dropbox.com/
Evernote	60MB upload max/month	Yes	Yes	convenient	terms/security questionable	http://www.evernote.com/
Google Docs/Drive (not via UConn)	15	Yes	Yes	convenient	terms/security questionable	http://drive.google.com
(Apple) iCloud	5	photos +iOS	photos +iOS	convenient	terms/security questionable	http://www.icloud.com
(MS) onedrive.com (was "skydrive")	7	Yes	Yes	convenient	terms/security questionable	http://www.skydrive.com
IV. Coming to UConn:						
Library-managed Repository*	TBD	Yes	Yes	secure, managed, long-term	not free above allotment; for archiving static data	http://classguides.lib.uconn.edu/datamanagement
<i>*available 2014</i>						

Active Data Working Space Decisions

Acronyms:

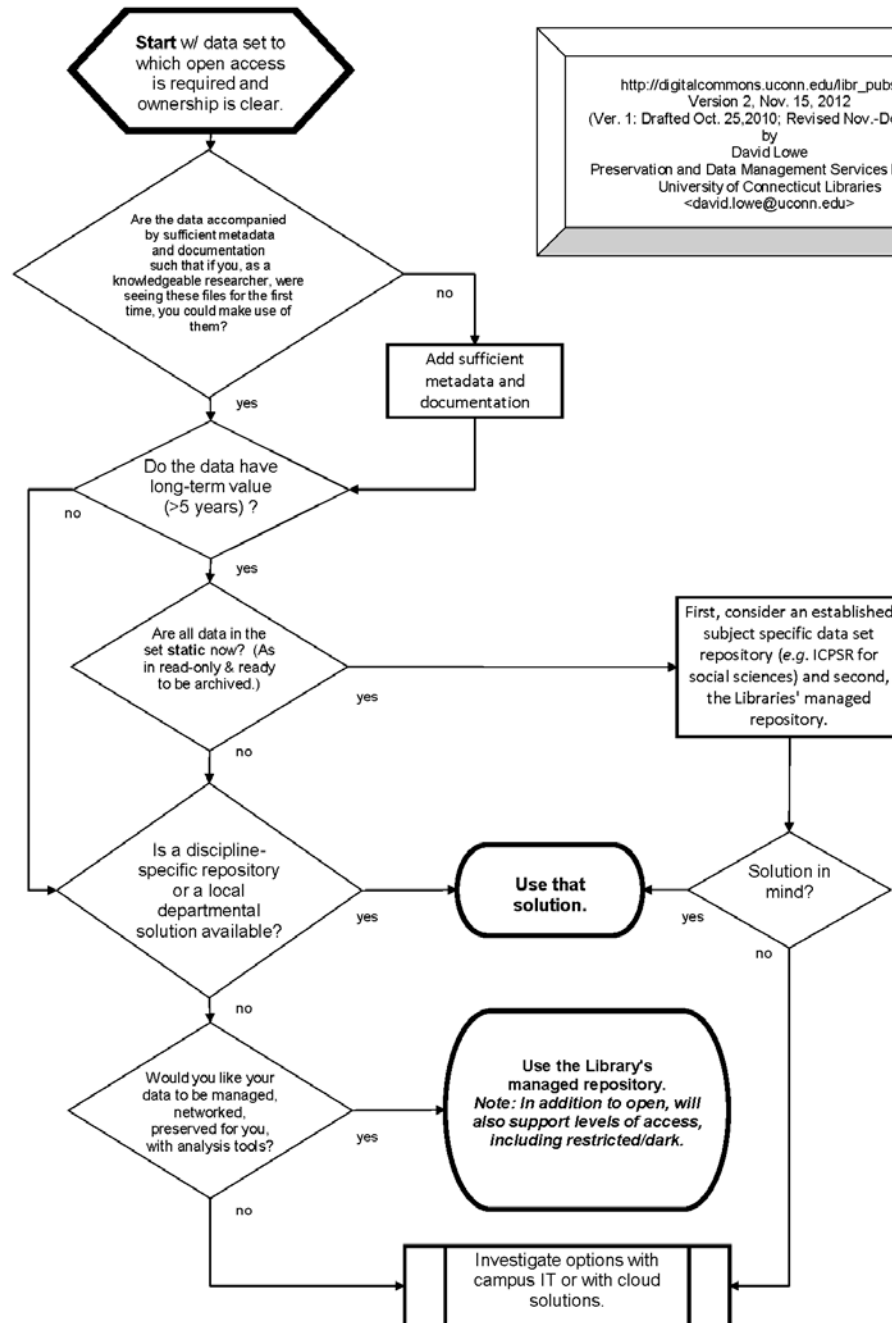
1. B2B (from options matrix):
Business-to-Business, *viz.*: UITS to your department
2. B2C (from options matrix):
Business-to-Customer, *viz.*: UITS to you
3. DoD: Department of Defense (restricted for national security)
4. ECM (from options matrix):
Enterprise Content Management
5. EFS (from options matrix):
Enterprise File Servers
6. NDA: Non-Disclosure Agreement (restricted by contract with a business entity)
7. PII: Personally Identifiable Information (restricted by laws for individual privacy)
8. RNAS: Research Networked Attached Storage

Determining Optimal Working Space for Academic Research Data in UConn's Environment, or:
"Where should I work with my files right now?"



Static Data Decisions

http://digitalcommons.uconn.edu/libr_pubs/33
 Version 2, Nov. 15, 2012
 (Ver. 1: Drafted Oct. 25, 2010; Revised Nov.-Dec. 2010)
 by
 David Lowe
 Preservation and Data Management Services Librarian
 University of Connecticut Libraries
 <david.lowe@uconn.edu>



Tightening Compliance Evidence from NSF Review Panel

- Fresh warning/cautionary tale from March 2014!
- PI had written in DMP (without our input):
 - “The data will be kept on the PI’s data cluster for ten years beyond the life of the project. The long-term preservation of digital objects ... is the PI’s ...responsibility.”
- Reviewer responded:
 - “NSF specifies that it is the University, not the PI, that has ultimate responsibility for long-term preservation of the data.”
 - Also questions access for research community

Near Future Work

- More decision trees!
 - Internal workflows
 - Embargoing and licensing levels
 - Cost modeling
- Direct costing institutional infrastructure
 - Within OMB Circular A-21 stipulations
 - Also per [NSF DMP FAQ #5](#):
 - does not have to be “unfunded mandate”

?

dave@uconn.edu

@davidblowe

s.uconn.edu/dave