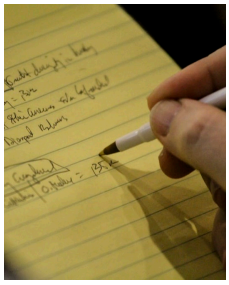
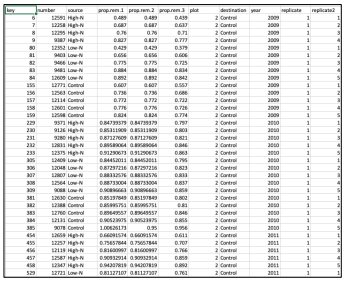
1

First some terminology…

Original Data



Operation Data



key	number	source	prop.rem.1	prop.rem.2	prop.rem.3	plot	destination	year	replicate	replicate2
6	12591	High-N	0.489	0.489	0.439	2	Control	2009	1	1
7	12258	High-N	0.687	0.687	0.637	2	Control	2009	1	2
8	12295	High-N	0.76	0.76	0.71	2	Control	2009	1	3
9	9387	High-N	0.827	0.827	0.777	2	Control	2009	1	4
80	12352	Low-N	0.423	0.423	0.373	2	Control	2009	1	1
81	9403	Low-N	0.656	0.656	0.606	2	Control	2009	1	2
82	9466	Low-N	0.775	0.775	0.725	2	Control	2009	1	3
83	9481	Low-N	0.884	0.884	0.834	2	Control	2009	1	4
84	12609	Low-N	0.892	0.892	0.842	2	Control	2009	1	5
155	12771	Control	0.607	0.607	0.557	2	Control	2009	1	1
156	12563	Control	0.736	0.736	0.686	2	Control	2009	1	2
157	12114	Control	0.772	0.772	0.722	2	Control	2009	1	3
158	12601	Control	0.776	0.776	0.726	2	Control	2009	1	4
159	12508	Control	0.824	0.824	0.774	2	Control	2009	1	5
229	9371	High-N	0.84739379	0.84739379	0.797	2	Control	2010	1	1
230	9126	High-N	0.85311909	0.85311909	0.803	2	Control	2010	1	2
231	9280	High-N	0.87127609	0.87127609	0.821	2	Control	2010	1	3
232	12831	High-N	0.89589064	0.89589064	0.846	2	Control	2010	1	4
233	12375	High-N	0.91290673	0.91290673	0.863	2	Control	2010	1	5
305	12409	Low-N	0.84452011	0.84452011	0.795	2	Control	2010	1	1
306	12048	Low-N	0.87297216	0.87297216	0.823	2	Control	2010	1	2
307	12807	Low-N	0.88332576	0.88332576	0.833	2	Control	2010	1	3
308	12564	Low-N	0.88733654	0.88733654	0.837	2	Control	2010	1	4
309	9088	Low-N	0.90896663	0.90896663	0.859	2	Control	2010	1	5
381	12630	Control	0.85197849	0.85197849	0.802	2	Control	2010	1	1
382	12388	Control	0.85995751	0.85995751	0.81	2	Control	2010	1	2
383	12760	Control	0.89649557	0.89649557	0.846	2	Control	2010	1	3
384	12131	Control	0.90523975	0.90523975	0.855	2	Control	2010	1	4
385	9078	Control	1.00626173	0.95	0.956	2	Control	2010	1	5
454	12659	High-N	0.66091574	0.66091574	0.611	2	Control	2011	1	1
455	12257	High-N	0.75657844	0.75657844	0.707	2	Control	2011	1	2
456	12119	High-N	0.81600997	0.81600997	0.766	2	Control	2011	1	3
457	12587	High-N	0.90032914	0.90032914	0.859	2	Control	2011	1	4
458	12347	High-N	0.94207819	0.94207819	0.892	2	Control	2011	1	5
529	12721	Low-N	0.81127107	0.81127107	0.761	2	Control	2011	1	1

Backup Data

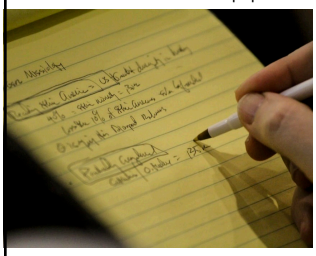
Data Archive

2

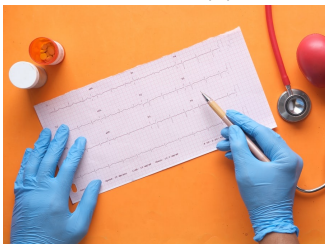
Original Data (i.e., 'Raw Data')

Information collected during research (plus the samples from which the info was extracted or the items on which you recorded it)

Information recorded on paper



Printouts from lab equipment

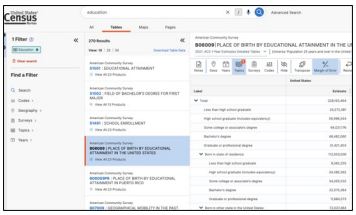


Photos

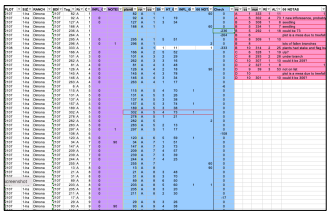
Audio recordings

Video recordings

Information from physical archives



Information from online archives



The spreadsheet into which you enter the information from any of these

3

Operation Data (i.e., 'Clean Data')

Original data after it has been Processed, Edited, Corrected, or Transformed.

Spreadsheet **after** proofreading and correcting values

key	number	source	prop.rem.1	prop.rem.2	prop.rem.3	plot	destination	year
6	12591	High-N	0.489	0.489	0.439	2	Control	2009
7	12258	High-N	0.687	0.687	0.637	2	Control	2009
8	12295	High-N	0.76	0.76	0.71	2	Control	2009
9	9387	High-N	0.827	0.827	0.777	2	Control	2009
80	12352	Low-N	0.429	0.429	0.379	2	Control	2009
81	9403	Low-N	0.656	0.656	0.606	2	Control	2009
82	9466	Low-N	0.775	0.775	0.725	2	Control	2009
83	9481	Low-N	0.884	0.884	0.834	2	Control	2009
84	12609	Low-N	0.892	0.892	0.842	2	Control	2009
155	12771	Control	0.607	0.607	0.557	2	Control	2009
156	12563	Control	0.736	0.736	0.686	2	Control	2009
157	12114	Control	0.772	0.772	0.722	2	Control	2009
158	12601	Control	0.776	0.776	0.726	2	Control	2009
159	12598	Control	0.824	0.824	0.774	2	Control	2009
229	9371	High-N	0.84739379	0.84739379	0.797	2	Control	2010
230	9126	High-N	0.85311909	0.85311909	0.803	2	Control	2010
231	9280	High-N	0.87127609	0.87127609	0.821	2	Control	2010
232	12831	High-N	0.89589064	0.89589064	0.846	2	Control	2010
233	12375	High-N	0.91290673	0.91290673	0.863	2	Control	2010
305	12409	Low-N	0.84452011	0.84452011	0.795	2	Control	2010
306	12048	Low-N	0.87297216	0.87297216	0.823	2	Control	2010
307	12807	Low-N	0.88332576	0.88332576	0.833	2	Control	2010
308	12564	Low-N	0.88733004	0.88733004	0.837	2	Control	2010
309	9088	Low-N	0.90896663	0.90896663	0.859	2	Control	2010
381	12630	Control	0.85197849	0.85197849	0.802	2	Control	2010
382	12388	Control	0.85995751	0.85995751	0.81	2	Control	2010
383	12760	Control	0.89649557	0.89649557	0.846	2	Control	2010

Edited and Corrected Transcripts

The standard Lorem Ipsum passage, used since the 1500s

"Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum."

Section 1.10.32 of "de Finibus Bonorum et Malorum", written by Cicero in 45 BC

"Sed ut perspiciatis unde omnis iste natus error sit voluptatem accusantium doloremque laudantium, totam rem aperiam, eaque ipsa quae ab illo inventore veritatis et quasi architecto beatae vitae dicta sunt explicabo. Nemo enim ipsam voluptatem quia voluptas sit aspernatur aut odit aut fugit, sed quia consequuntur magni dolores eos qui ratione voluptatem sequi nesciunt. Neque porro quisquam est, qui dolorem ipsum quia dolor sit amet, consectetur, adipisci velit, sed quia non numquam eius modi tempora incidunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim ad minima veniam, quis nostrum exercitationem ullam corporis suscipit laboriosam, nisi ut aliquid ex ea commodi consequatur? Quis autem vel eum iure reprehenderit qui in ea voluptate velit esse quam nihil molestiae consequatur, vel illum qui dolorem eum fugiat quo voluptas nulla pariatur?"

1914 translation by H. Rackham

"But I must explain to you how all this mistaken idea of denouncing pleasure and praising pain was born and I will give you a complete account of the system, and expound the actual teachings of the great explorer of the truth, the master-builder of human happiness. No one rejects, dislikes, or avoids pleasure itself, because it is pleasure, but because those who do not know how to pursue pleasure rationally encounter consequences that are extremely painful. Nor again is there anyone who loves or pursues or desires to obtain pain of itself, because it is pain, but because occasionally circumstances occur in which toil and pain can procure him some great pleasure. To take a trivial example, which of us ever undertakes laborious physical exercise, except to obtain some advantage from it? But who has any right to find fault with a man who chooses to enjoy a pleasure that has no annoying consequences, or one who avoids a pain that produces no resultant pleasure?"

Section 1.10.33 of "de Finibus Bonorum et Malorum", written by Cicero in 45 BC

"At vero eos et accusamus et iusto odio dignissimos ducimus qui blanditiis praesentium voluptatum deleniti atque corrupti quos dolores et quas molestias excepturi sint occaecati cupiditate non provident, similique sunt in culpa qui officia deserunt mollitia animi, id est laborum et dolorum fuga. Et harum quidem rerum facilis est et expedita distinctio. Nam libero tempore, cum soluta nobis est eligendi optio cumque nihil impedit quo minus id quod maxime placeat facere possimus, omnis voluptas assumenda est, omnis dolor repellendus. Temporibus autem quibusdam et aut officiis debitis aut rerum necessitatibus saepe eveniet ut et voluptates repudiandae sint et molestiae non recusandae. Itaque earum rerum hic tenetur a sapiente delectus, ut aut reiciendis voluptatibus maiores alias consequatur aut perferendis doloribus asperiores repellat."

These are the ones on which you carry out 'operations' , i.e., analyses.

4

Backup Data

Copies of 'Original' & 'Operation' Data allowing you to resume analyses if files are lost or damaged.

Electronic version stored on your computer, the cloud, or...

...a portable storage unit.

Paper Copies

5

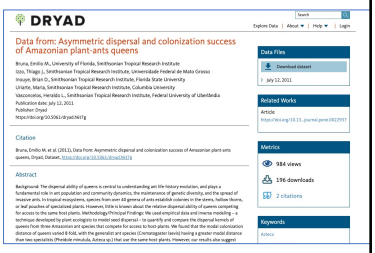
Data Archive

Long-term storage of Original & Operational Data in a format & location where it is :
(1) secure and (2) can be readily discovered and used by others.

Museum Collections

Library

Online Data Repository



6

Why bother backing up data?

Because bad things happen.

Especially to you.

7

8

Other common causes of data loss include…

9

DEMONIC INTERVENTION

natural disasters

pests

Cats

hackers

document loss

ACCIDENTAL DELETION

VIRUSES & MALWARE

software or media obsolescence

THEFT

file corruption

ACCIDENTS

human error

POWER FAILURE

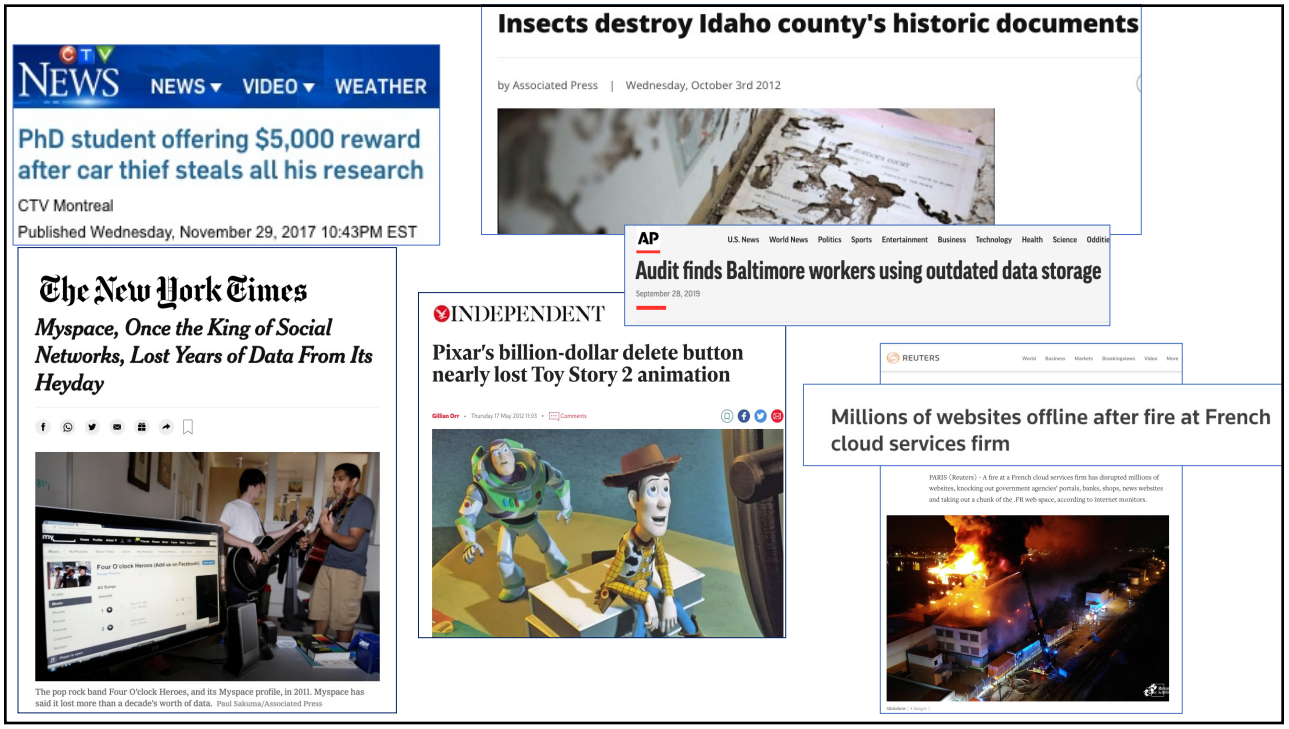
HARD DRIVE FAILURE

coffee spilled on laptop

10

These things happen more often and on bigger scales than you think.

11



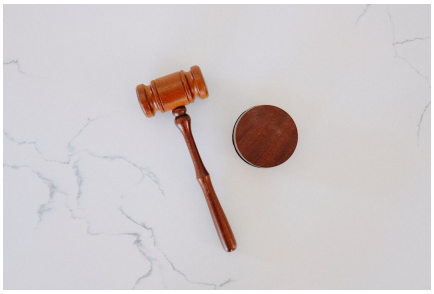
12

Why else should we bother backing up data?

Because you are required to.

13

Compliance with: Public Records Requests, lawsuits, funder mandates…



National Institutes of Health

And finally….

14

Because it makes research and collaboration easier, more efficient, and more productive.

15

OK…so what should we be doing to back up our data?

16

3-2-1

3 copies of your data stored on
2 different types of media, with
1 copy stored off-site

17

What does ‘off-site’ mean?

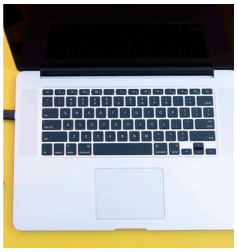
If you WORK in your office → home could be OFF-SITE

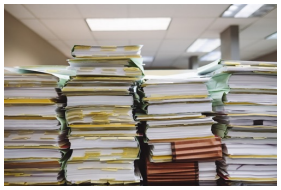
Then your office could be OFF-SITE ← If you WORK from home

18

"Different" Storage Media could be:

- Desktop Computer
- Laptop Computer
- External hard drive
- Campus server
- Commercial Cloud Backup
- Optical Storage (CD, DVD)
- Paper
  original notebooks & data sheets, printouts of spreadsheets, xerox copies of originals, etc.





19

Keep in mind that some media can become obsolete.....

20

While others do not.

21

WARNING!!

22

Synchronized cloud storage is *not* "**off-site**" or "**different**".

Why?

Deletion from Location 1 **automatically** triggers deletion from Location 2.

23

IMPORTANT: There are special considerations for "Restricted Data"

- **Find out** if your data require extra protection, and
- **Get guidance** from your institution on the additional steps and resources needed to protect original and backups of Restricted Data

24

And finally….

To ensure that your data can be read by as many computer programs as possible for many years into the future :

USE OPEN

(a.k.a. 'non-proprietary')

DATA FORMATS

25

And be sure to…

- Test Your Storage Media Regularly
- Beware of Early Hardware Failures
- Determine the Life of Your Hard Drives
- Routinely Inspect and Replace Data Discs
- Handle and Store Your Media With Care

26

Everything is going to be ok?

27

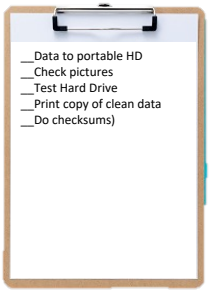
Make a plan. Write it down.

- what needs to be prepared and for how long
- where backups are located
- who can access backups and how contacted
- how often data should be backed up
- what kind of backups are performed
- who is responsible for performing the backups
- hardware and software used for performing backups
- how / how often to check if backups is successful
- the media are used to backup data
- a list of any data that are *not* archived or backed up

Then be sure to…

28

1. Make it easy to follow your plan



2. Automate as much of the process as you can.

3. Revisit & Revise your plan regularly

*And remember: synced computer and cloud drive are **not** independent copies*

29

The End

(off-site)

Photos: Unsplash or EM Bruna
Music: www.bensound.com

30

What about the The Originals?

Consider the volume and security needs – both short-term & long-term – of specimens, samples, documents and data sheets, photographs, or other physical items. Depending on the situation you might use:

- Campus Office
- Lab or Museum/Herbarium
- Departmental Office
- Commercial storage facility
- Home
- Library

31

Photo, video, and audio backup:

More challenging, in part because transcriptions & captioning can be important for interpretation, discovery, and accessibility.

Be sure to back-up both the original files *and* the transcriptions.

Think ahead: storage can also be challenging because the files are large and there are often large numbers of them.

32