Organizing and Safeguarding Information on Disk

M. E. Kabay, PhD, CISSP-ISSMP

Organizing and Safeguarding Information on Disk

M. E. Kabay, PhD, CISSP-ISSMP

Professor of Computer Information Systems
School of Business and Management, College of Professional Schools
Norwich University, Northfield, VT 05663-1035 USA
mailto:mkabay@norwich.edu

Contents

1   Introduction

Students and colleagues how I manage information that I use in teaching, research, writing articles and books, have sometimes asked me. A recent question is, "I am … curious how you manage data you come across. I sense that when you are writing a paper or book and need to tap into some past work you or someone else has done, you have organized your hard drive in such a way that you don't have to hunt too long to find it (maybe you use a database of some sort rather than a directory structure)."

This essay summarizes how I handle information on disk. I don't claim that my methods are ideal or based in any particular theory; I just don't like clutter, and I like to be able to find stuff when I need it. I also dislike having to redo my work, so I'll also summarize how I back up my data.

The figures in this document mostly show my systems in 2016, which use Windows 10. However, the guidelines can apply to any operating system with suitable adaptation. I hope that you will find at least some of the suggestions useful. Just be sure that you understand right from the start that my main point is that you can benefit by figuring out a system that suits *your* needs – not that my system should be adopted by everyone.

2 Disk Drives

First of all, it may be useful to have multiple disk drives or to partition disk drives into virtual drives with different functions. For example, on my systems, the drives at time of writing are as shown in Figure 1:
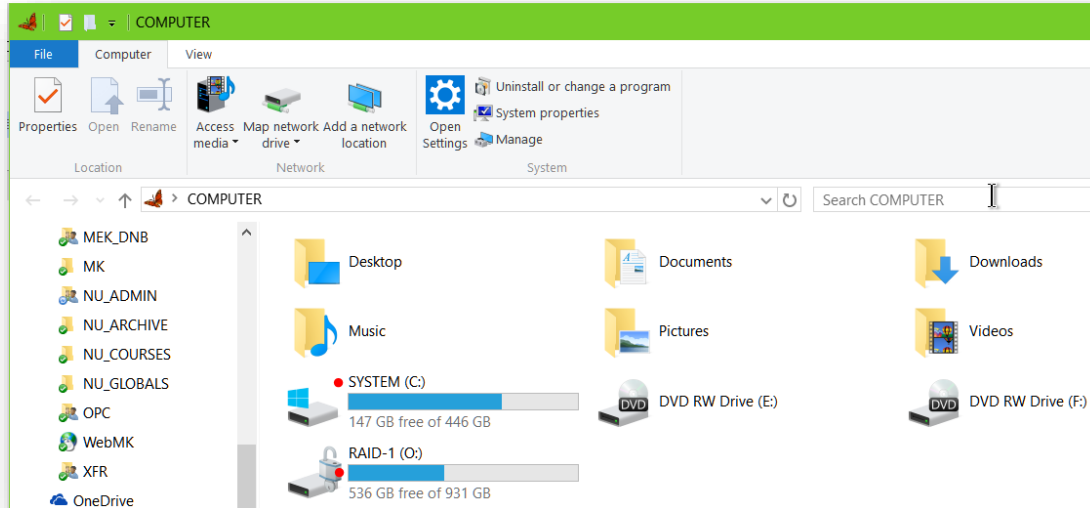

Figure 1. Hard disks.

Separate disk drives (or named partitions on a single drive) help to segregate data for performance and security functions. Keeping as much user data as possible off the operating-system (OS) disk (here, the solid-state disk C:) maximizes throughput on the disk most used by the operating system. Putting other data on the RAID 1[1] mirrored drive O: also makes it easier to handle backup/restore operations. The O: drive shown above is also encrypted using the Windows BitLocker tool (**Control Panel\All Control Panel Items\BitLocker Drive Encryption**). Encryption means that even if the drive were stolen (or discarded without wiping it), the data would be difficult (practically speaking, impossible for anyone but government agencies or expert forensic scientists) to decrypt.

Figure 2 shows the visible folders on the system drive. As you can see, I hide folders that I rarely access, leaving only two visible (**Dropbox** and **itunes**). The **Show/hide** section of the Windows 10 ribbon includes a checkbox for showing hidden items at will. See Figure 3.
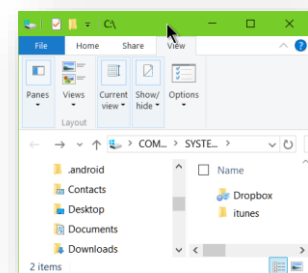

Figure 2. Visible folders on C drive.

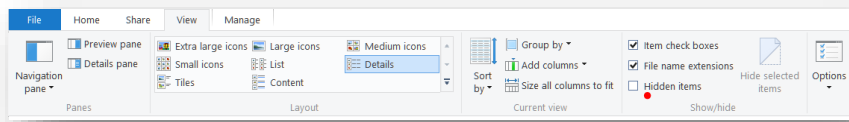**Figure 3. Show/hide options on Windows 10 ribbon.**

---

[1] RAID 1 (Redundant Array of Independent Disks Mode 1) combines two physical disk drives into a single logical drive in which everything written to one drive is almost instantly duplicated on the other drive. If the probability of failure of each of the drives is p, then the probability that both will fail at the same time is p^2 – much smaller.

3   Active User Files

All my active data are stored in a **Dropbox** cloud-storage folder (see Figure 4. Folders in Dropbox.Figure 4). The data are encrypted in the cloud (storage services on remote servers accessed through the Internet). As you can see from, the folder structure of my regularly-used data is quite simple.

Figure 4. Folders in Dropbox.

- BU contains utility files for my MAIN, SPARE and LAPTOP computers.
- **CYBERSAFETY** is a shared folder for a writing project in which contributors can access the files.
- **DB** is a repository for active reference materials such as OUTLOOK files, maps, TODO database and so on.
- **MEK_DNB** is a shared folder for exchanging files easily with my wife, Dr Deborah N. Black, MDCM, FRCP(C)
- **MK** contains a variety of personal information such as financial records
- **NU_ADMIN** has files for my work at the University
- **NU_ARCHIVE** has all the teaching materials and student records starting with my first courses in Fall 2001
- **NU_COURSES** has teaching-related materials
- **NU_GLOBALS** has reference materials from the University
- **OPC** is a shared folder for one of my clients
- **WebMK** has the source code for all the pages of my Website, mekabay.com
- **XFR** is used for transferring files quickly among my three computers. Typically, screenshots and reports are stored there while they are useful (e.g., for attaching to email messages) and then moved to an archive.
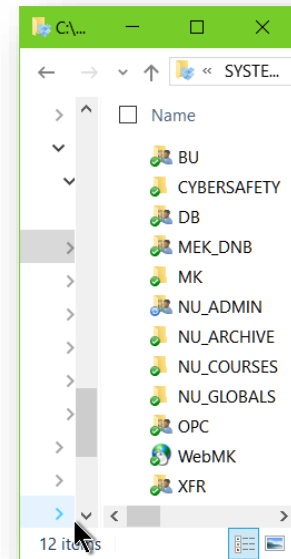
3.1   Folders in BU

Figure 5 shows folders in the **BU** folder of **Dropbox**. These locations allow different computers sharing access to **Dropbox** to standardize such issues as

- Digital signature for signing Adobe Acrobat PDF files
- Backups of antivirus/antispam programs
- Configurations for backup programs
- Accounting backups
- Settings for macro facilities
- System settings for each system (main, spare, laptop)
  - o   Registry backups
  - o   Screen icon layouts for left and right screens
  - o   Bitlocker recovery key where appropriate
- Customization files (*_Customizations.exportedUI) for MS Office program interfaces (ribbon layout, Quick Access Toolbar) to permit synchronization of the user interface across systems
- Archives of files moved out of the My Documents and XFR folders
- Synchronization data for my preferred browser (Opera)
- Global PGP encryption keys
- Skype contact backups
- Microsoft Windows system encryption keys
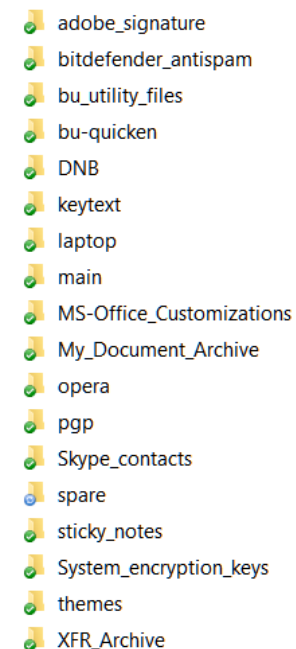- Icon themes saved from Windows 95 (!)

Figure 5. Contents of BU folder.

3.2   Information in Personal (MK) Folder

The folder where I keep current personal information (Figure 6) includes

- The current publication I'm working on
- Templates and examples of cards (birthday, business, seasonal), forms and labels
- Correspondence, reports, and files for consulting clients
- Personal correspondence
- Business correspondence
- Curriculum vitae (resumé)
- Family documents and photographs
- Financial and legal records
- Mentoring projects (helping specific students improve their academic success)
- Miscellaneous documents not classified elsewhere
- Excel templates for various functions

Figure 6. Information in personal folder (MK).

3.3   Organizing Financial Information

Since much of our financial information is important for yearly tax returns, I put each year's information in its own folder (Figure 7). The current year's information is in folder **_current**, which sorts to the top when using a normal alphabetic sort for the folder names. At the end of each year, I copy **_current** to the appropriate year-folder and then empty the files out of **_current**.
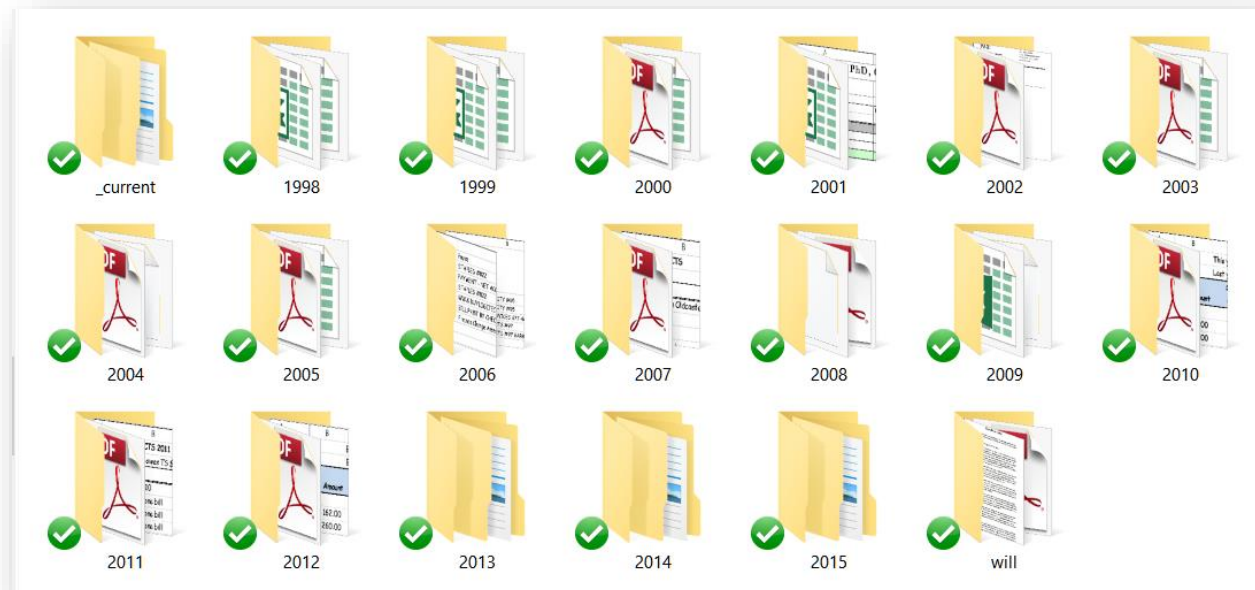

Figure 7. Inside the Financial & Legal folder.

4   Work-related Documents

4.1   Administrative Work

NU_ADMIN has information related to my work at Norwich University; **NU_ARCHIVE** has historical records for all my courses classified by semester (Figure 8).

Inside the **b&m** folder are folders for specific aspects of my work as a professor in the School of Business and Management(Figure 9).

Naturally, the hierarchy goes as far down as necessary, so several of these folders have their own folders inside them. Readers can simply extend the principle without further elaboration here.
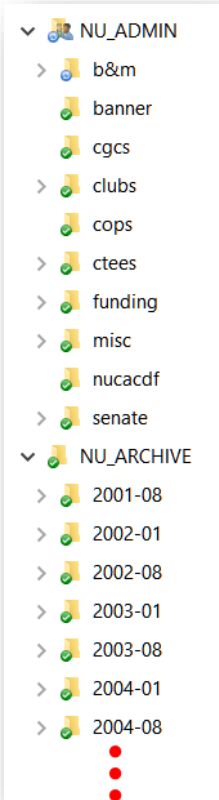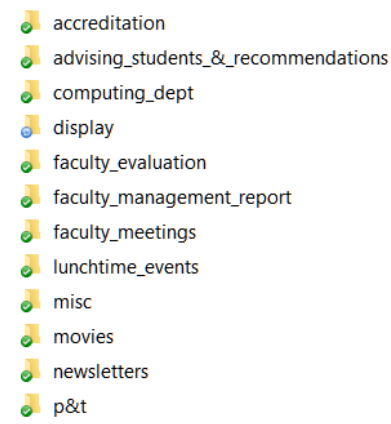

Figure 8. Administrative Information & Start of Course Archives.

Figure 9. Details of the Folder for Work Specific to the School of Business & Management.

4.2 Current and Future Courses

My work requires access to all the teaching materials and administrative functions for the courses I'm teaching in the current semester and also a place to prepare materials for courses coming up in the next semester. Figure 10 shows the contents of the **NU_COURSES** folder.
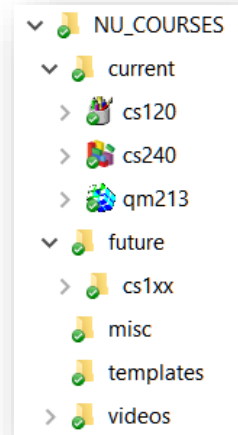

Figure 10. Course-related folders.

Current course folders have similar structures, reflecting the usual ways I provide teaching materials and keep records. Figure 11 shows two courses in the **current** folder for Spring 2016. Note the similarities and differences. I added an **old** folder for the **cs240** (database) class because I wanted quick access to materials I used the last time I taught database (2010). After I finish with those materials, I'll delete the **old** folder.

4.3 Globals

The **NU_GLOBALS** folder has useful documents, as shown in Figure 12.

4.4 The WebMK Folder

The **WebMK** folder is the source for all the material posted to my Website, mekabay.com (from which you presumably downloaded this file). In addition to all the folders found on the Website, it includes source files for the PDF documents I make available as well as the materials for folders that are part of the Dark Web: they have never had any pointers to them posted online, so they are invisible to search engines such as GOOGLE or BING. In addition, should my routine check find that someone I allowed into the personal folders has posted a searchable link to that section of my Website, I simply rename it to a different random sequence.

Figure 11. Typical structure of course folders.

Figure 12. The Globals folder.
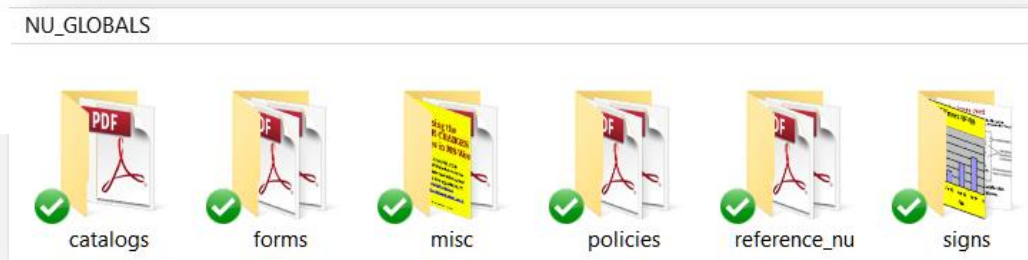
5   Data Folder

The **Data** folder (Figure 13) has mostly (but not exclusively) archival material including

- Correspondence and documents relating to various professional and volunteer associations for which I have volunteered
- Conference submissions and proceedings (but not teaching materials)
- Personal photo albums
- Publications (including whatever I'm working on currently)
- Non-Norwich teaching materials (e.g., courses for other colleges, for invited lectures, and for conferences)
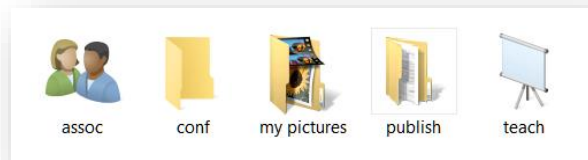
Figure 13. The Data folder.

6   Reference Folder

The **Reference** folder (Figure 14) has largely archival material that is occasionally useful. I store

- Roughly 10,000 white papers and other documentation about specific topics in my field in the archives folder (often used to help students in their projects)
- More than 100,000 licensed images in the clipart folder (some dating back to a multi-CD set bought in 1995!)
- About 50,000 files from electronic books and training programs in the ebooks folder
- Installation files for software used on each of my three computers plus their license documents
- Some 11,000 files for training videos used in my courses

Figure 14. The Reference folder.

7 Backing Up

7.1 Synchronizing Files Across Systems and Removable Disks

Every day, I copy all changes on my active systems (MAIN or SPARE) to one of two 1 TB removable disks (called Q: on each of the two computers). The next day, I copy all the day's changes to the other removable disk. In addition, all my active files are stored on Dropbox, so both MAIN and SPARE and also LAPTOP have identic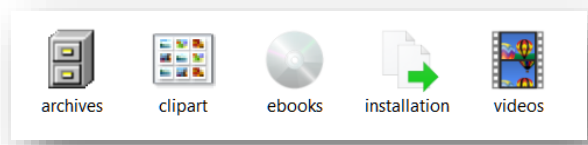al contents within minutes of each other if they are online. I typically keep one of the two removable disks at my Norwich University office and the other at my home office. That way, if there were a disaster or a theft at one of the offices that resulted in disappearance or destruction of one of the drives, the other would provide a backup no older than a couple of days.

7.2 Backing Up the Day's Changes

An *incremental backup* stores all the files that have changed since the last backup of any type. I perform an incremental backup using the WinZIP compression program at least once a week – often every day – and store the compressed files in a folder at **O:\bu_incremental_current**. At the end of each month, I move all the month's incremental backups into a folder at **O:\bu_archives**. Each month has its own folder there; see Figure 15 for an example of a month's worth of ZIPX files.
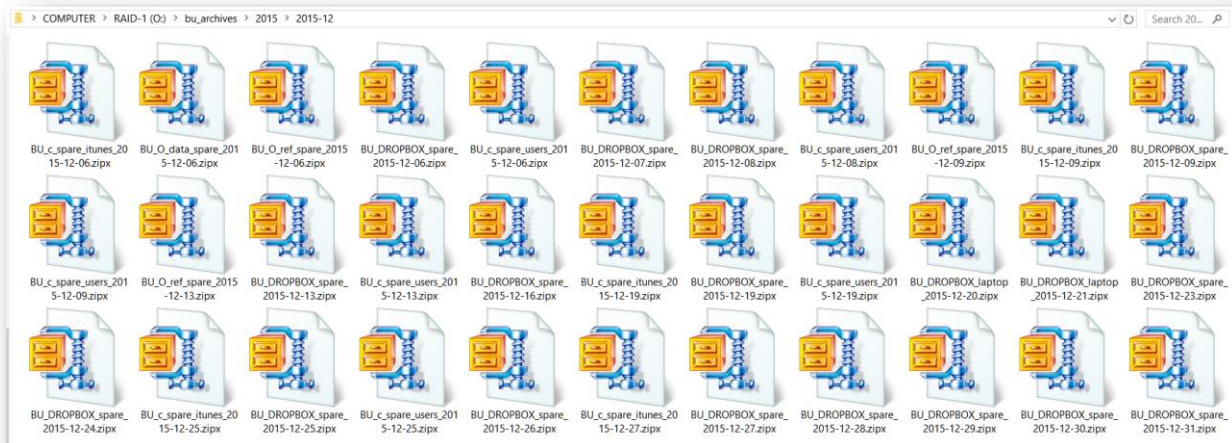
Figure 15. Incremental backup ZIPX files for December 2015.

7.3 Monthly Full Backups

At the end of every month, I move all the incremental backups into a folder labeled with the month (as shown above in Figure 15). Then I send copies of those incremental backups to an external hard drive that is kept in a fire-resistant safe (Figure 16) in my home office. The safe is rated at three hours of resistance to a fire; it has a three-position combination and I keep the dial at zero when I lock it (to eliminate all knowledge of probable elements of the combination).

After storing the incrementals, I store all the data from drives C: and O: on a backup drive.

Figure 16. Office safe showing backup drives.

To keep track of all the backups and on which drives they are stored, I keep a simple list (Figure 18) in Excel with sorts by date, by drive and by type plus a page showing freespace on each drive (Figure 17).

	DATE	DRIVE	TYPE	Notes		DRIVE	DATE	TYPE	Notes		TYPE	DATE	DRIVE	Notes
1	**BACKUP**	**as of**		**2016-01-05**										
2	*By Date*					*By Drive*					*By Content*			
3	DATE	DRIVE	TYPE	Notes		DRIVE	DATE	TYPE	Notes		TYPE	DATE	DRIVE	Notes
15	2014-09-30	BU1B	P			BU1B	2014-11-30	P			F	2015-06-01	BU8	O ONLY
16	2014-09-30	BU7	F			BU1B	2014-12-31	P			F	2015-07-01	BU7	
17	2014-10-31	BU1B	P			BU1B	2015-01-31	P			F	2015-08-01	BU4	
18	2014-11-30	BU1B	P			BU1B	2015-04-30	P			F	2015-11-01	BU4	Dropbox & iTunes
19	2014-12-01	BU7	F			BU1B	2015-06-01	P			F	2015-12-31	BU6	C, O, image
20	2014-12-31	BU1B	P			BU1B	2015-07-01	P			I	2015-05-02	BU3B	MAIN IMAGE
21	2015-01-01	BU2	F			BU1B	2015-07-31	P			P	2014-01-31	BU1B	
22	2015-01-31	BU1A	F			BU1B	2015-10-03	P			P	2014-02-28	BU1B	
23	2015-01-31	BU1B	P			BU1B	2015-11-01	P			P	2014-03-31	BU1B	
24	2015-04-01	BU3A	F	PGD		BU1B	2015-12-02	P			P	2014-04-30	BU1B	
25	2015-04-30	BU1B	P			BU1B	2015-12-31	P			P	2014-05-31	BU1B	
26	2015-05-02	BU3A	F	O		BU2	2015-01-01	F			P	2014-06-30	BU1B	
27	2015-05-02	BU3B	I	MAIN IMAGE		BU2	2015-06-01	F	C ONLY		P	2014-07-31	BU1B	
28	2015-05-03	BU3A	F	iTunes		BU3A	2015-04-01	F	PGD		P	2014-08-31	BU1B	
29	2015-06-01	BU1B	P			BU3A	2015-05-02	F	O		P	2014-09-30	BU1B	
30	2015-06-01	BU2	F	C ONLY		BU3A	2015-05-03	F	iTunes		P	2014-10-31	BU1B	
31	2015-06-01	BU8	F	O ONLY		BU3B	2015-05-02	I	MAIN IMAGE		P	2014-11-30	BU1B	
32	2015-07-01	BU1B	P			BU4	2015-08-01	F			P	2014-12-31	BU1B	
33	2015-07-01	BU7	F			BU4	2015-11-01	F	Dropbox & iTunes		P	2015-01-31	BU1B	
34	2015-07-31	BU1B	P			BU6	2015-12-31	F	C, O, image		P	2015-04-30	BU1B	
35	2015-08-01	BU4	F			BU7	2014-12-01	F			P	2015-06-01	BU1B	
36	2015-10-03	BU1B	P			BU7	2015-07-01	F			P	2015-07-01	BU1B	
37	2015-11-01	BU1B	P			BU7	2014-09-30	F			P	2015-07-31	BU1B	
38	2015-11-01	BU4	F	Dropbox & iTunes		BU8	2014-06-01	F	C-data & PGD		P	2015-10-03	BU1B	
39	2015-12-02	BU1B	P			BU8	2015-06-01	F	O ONLY		P	2015-11-01	BU1B	
40	2015-12-31	BU1B	P			BU9	2014-04-30	F	PGD files		P	2015-12-02	BU1B	
41	2015-12-31	BU6	F	C, O, image		BU9	2014-06-30	F	Data only -- no PGD		P	2015-12-31	BU1B	

Figure 18. Backup list.

	DRIVE	Capacity (GB)	FREE (GB)	Freespace
1	**DRIVE**	**Capacity (GB)**	**FREE (GB)**	**Freespace**
2	1A	465	34	7%
3	1B*	465	177	38%
4	2	931	21	2%
5	3A	465	42.6	9%
6	3B	465	107	23%
7	4	930	14	2%
8	6	1864	1485	80%
9	7	1393	98	7%
10	8	465	42	9%
11	9	465	465	100%
12	**TOTAL**	**7908**	**2485.6**	**31%**
13	*PGD FILES TAKE UP ~71 GB*			
14	*Use this one for partials*			

Figure 17. Freespace chart.

Observant readers will note that there is no drive 5 – it was a 2 TB drive that went bad. When I replace it, the new drive will be labeled as 5.

I keep "partials" (P) – an old word that is ambiguous because it could refer to incrementals or to *differentials* (backups of everything that's changed since the last full backup) on a single drive for convenience.

7.4   Quarterly Image Backups

About every three months, I use special software to create an *image* backup that would permit restoration of the entire configuration of one of my two personal computers – MAIN in one case, SPARE in the next. Readers can find a wide range of programs to evaluate using a search with keywords < windows 10 image software > or whatever operating system is appropriate.

Under Windows 10, it is possible that you may be able to create an image without buying any other software, but that service may be incompatible with your hardware (as it is for me). Start in the Control Panel and access the **File History** menu.
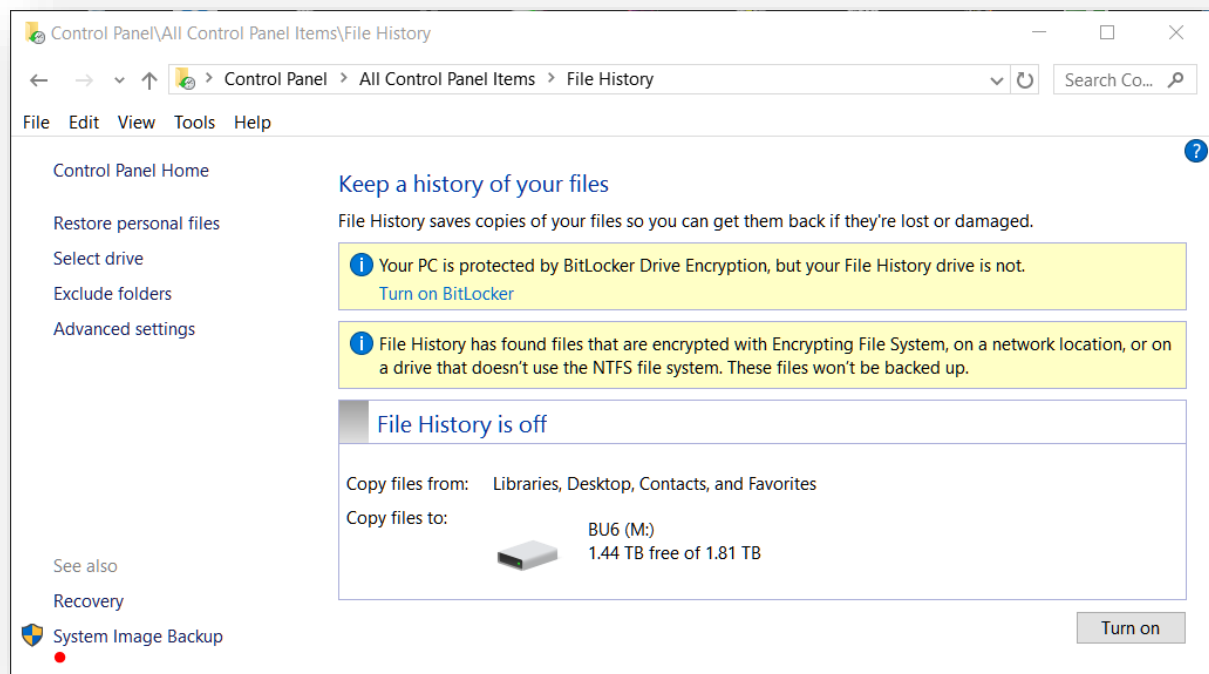
Select the **System Image Backup**.

Figure 19. Accessing System Image Backup using File History Menu.

In the next menu (Figure 20), select **Create a system image** and continue through the easy sequence of choices to put one on a selected drive.

I have never succeeded in completing a backup with this tool, so I use a commercial product called **Laplink DiskImage Professional Edition**. This information should not be construed as a recommendation.
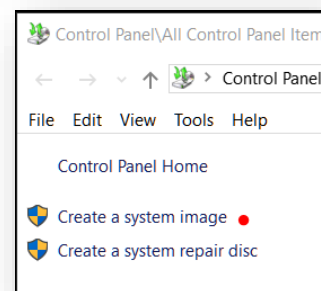

Figure 20. Backup and Restore Options.

8   Versions of Projects

One of the valuable practices I recommend to all my students is to keep versions of their project files. For example, shows a couple of versions of a manuscript I worked on in the summer of 2015.
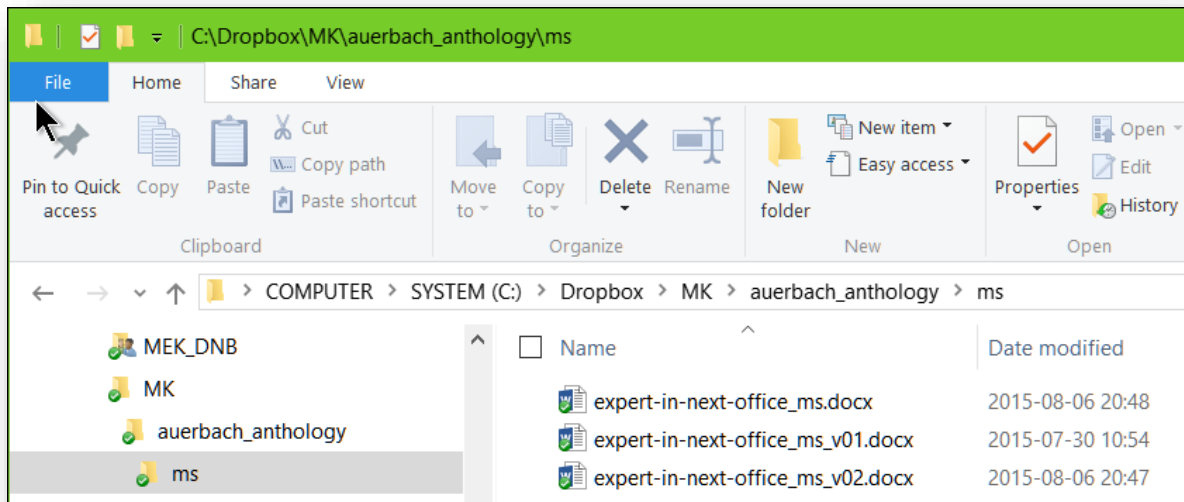

Figure 21. Versions of a Manuscript.

As you can see, at the end of a workday (in this case, August 6), I saved the file as **expert-in-the-next-office_ms_v02.docx** and then as what is called the *canonical* version (think of it as the most recent version), **expert-in-the-next-office_ms.docx** – this latter being the one I always open when starting to work on the file. The next time I work on the file, I will save it as **expert-in-the-next-office_ms_v03.docx** and continue in that fashion.

The benefits of this practice are several:

- If one accidentally damages or deletes a portion of the text, the previous version(s) may provide access to the undamaged material.
- If one is charged with plagiarism because someone else steals the file and submits it for credit first, it is possible that the backups of the dated, sequential version files can prove that the original author is actually the author and not the thief.
- If a system crashes hard, the external backups may save a great deal of time in recovering from the disaster so that the project can be completed on time.

9   Concluding Remarks

I hope that you have found this guided tour useful. Remember, the important thing is not that you slavishly follow someone else's recommendations – it's that you think about how you use your tools and make rational decisions about optimizing your patterns of use.

Good luck!