

Professors Emeriti Network

Professors Emeriti Network. We walk what we talk.

How Retired Professors are Organized and how they eCollaborate
Zoom Meeting on Wednesday, November 19, 2025, 17:00-19:00 CET

Getting Organised

Dr. Heinz V Dreher, Independent Researcher, Professor Emeritus of Informatics
Mahogany Creek, Western Australia
textminer@outlook.com

My goals -

1. all information containing objects (documents) created in or converted to digital form, mainly into pdf files;
2. progressively convert previously created or occurring documents into digital form;
3. all objects to be searchable at the meta level (i.e. via metadata) and preferably via the content itself;
4. all objects to be available across various devices and locations (i.e. use cloud storage) and where practicable and safe;
5. find easy to use, functional, and highly reliable Apps - preferably just one App that 'does it all' - (probably a 'tall order');
6. chosen Apps must interoperate so that App malfunction or failure is not catastrophic;
7. chosen digital object format need to be persistent (not subject to vagaries or fads)

and then one can *mine the text* to "one's heart's desire"

Now, two examples of goal achievement:

- one for <financial documents>, and - one for <published articles>

Artefact Acquisition, pre-processing, and filing (storage) <financial documents> - mainly purchase receipts, and usually in a paper-based format —> my procedure relies on the Notes-App (native to all Apple devices), and a purchased App named DEVONthink 4, from devontechnologies.com - i have a paid license for desktop, laptop, and tablet devices.

** step 1 - using iPhone Notes-app, photograph the object so that there is just the receipt and no background {use \Scan DocumentsX function which will automatically create an object in the chosen folder of your Notes-app}.

** step 2 - using the DT-app (short for whichever version of DevonThink i am using), import a copy of the object created in step 1.

** apply the OCR function if it has not been done automatically and ensure that the object is saved in my DT-database named appropriately, in my case 9-receipts (i have organised all my stuff into 16 partitions and each one has its own DT-database).

<published articles> - these are scientific articles, published reports, and documents that i have acquired and want good access to (where good = immediate, anywhere, anytime, and searchable at both metadata level and content level).

** since these artefacts are already in digital format i only need to be concerned with them being in \PDF+text\ format which is mostly realised automatically by applying ** step 2, from above.

Outcomes or Results

Filename is <mdg coles coffee food \$103,30.jpg> size 1.5MB

```
Store: 330 - CS MJICARIIIc Store Manager: Kelly Ftlorrrf; 08 0/98 69LMI
ServoH By: Tatjow
Register: 003
Date: 22/09/2025
Receipt: 6701
TIN: 15.15
Description
1 1WI RECYCLE PAPER b UACH          >
      21 $0.25 EACH                    0.50
MEHTTA COFFEE BLUE 25UÄAM          80.00
      101 $2.00 EACH
CIS FR CWN BMCSIIC PERKG           7.74
BAHAMAS PERKG                       6.06
      1.346 kg e1 - $4.50/kg
BLUEBERRIES 125GRAM                9.00
      2 - $4.50 EACH
Total for 16 Itees:                 $103.30
EFT                                   $103.X
GST INCLUDED IN TOTAL                $0.05
Coles 22/09/25 15:15 ***** 8102 CREDIT ACCOUNT APSK 0001 AFC 0132 PLRCMASE RR* XM03<-70100 AUTH 07154*
MA MJ N33003 MASIERCARD Debit MasterCard 400000.041010 AUDI 103.30 (00)A>WOVED
M721366
```

It's somewhat tedious to change the autoassigned filename but there are still advantages to this practice and I have been doing it for many years. At some point I will figure out a better way, as I have done with my digital photos, of which there are many hundreds being acquired at a time.

The OCR procedure in the DT-app is very powerful and it creates, in this case a file with name:

<mdg coles coffee food \$103,30.pdf> and size 122KB.

Note the 'order of magnitude reduction' in the item size. But it's the ability to find text within the PDF that is so useful. Being mindful of my wish to have access to all important/useful

documents remotely (whilst on-the-road as it were) I have arranged my data into 16 databases (standalone repositories of information) and assigned the following names:

n 1a_email messages.dtBase2	43.2 MB
B 1b-HvD health.dtBase2	1.03 GB
B 1c-HvD projects.dtBase2	2.58 GB
B 2a-all year-end stories.dtBase2	2.9 GB
B 2b-4WD trips.dtBase2	1.18 GB
B 3a-HvD devices - apps.dtBase2	3.55 GB
B 3b-HvD solar - electrical stuff.dtBase2	1.95 GB
B 3c-MCK buildings - land plans.dtBase2	22.3 MB
B 3d-HvD vehicles.dtBase2	803.9 MB
B 4-search subsets.dtBase2	1.03 GB
B 5-natural systems.dtBase2	869.3 MB
B 6-genealogies.dtBase2	25.1 MB
B 7-published articles.dtBase2	2.27 GB
B 8a-mdg shire stuff.dtBase2	7.17 GB
B 8b-noise issues.dtBase2	2.32 GB
B 9- receipts.dtBase2	2.48 GB

Empowering functions

-
- Say i need/want to discover the number of times i purchased coffee in recent years - using the DT-database with name <9-receipts> into which i put all my financial transactions, i search for “coffee” figuring that this may reveal some useful information (58 purchases).
- But i also see that a couple of results are inexplicable, reminding me that nothing is perfect.
-
- Or i may need/want to see something about my blueberry purchases - frequency, cost, etc., i set the scope of a search in DT to the <9-receipts> database and enter the search string “blueberr” as this will find the words “blueberry” and “blueberries”, and so on.

*** do a live demo of this ***

-
- In discussion with a friend, the topic of solar panels came up - what panels did i select? Using the <3b-HvDsolar & electrical stuff> database i search for the term “solar panel” and quickly find 116 items (documents) that may pertain, and because i assign a ‘meaningful’ filename to most items i put in the DT-databases, i can scan (using my eyes) the search results list and select an item that seems to fit. There i can see what i wanted and of course lots more information besides.

In a discussion around one of by outback campfires, about the local indigenous people and their culture, the question of cannibalism arose. As i am interested in the affairs of the Australian Aboriginals, and have read widely on this topic, there exist a few hundred articles to be searched. Very quickly i was able to find some authoritative material to inform the discussion - even in the absence of internet connectivity.