

# DPC Student Conference

## Things I wish I knew before I started

January, 2015

Dave Thompson  
Digital Curator, Wellcome Library

Good morning. Everything is fine. Please do this. Everything is fine. Unexpected event. It works. Please change it. Everything is fine. Why does this happen? We've lost internet access. Take a deep breath. This book for the launch isn't digitised. Everything is fine. Everything is fine. Can you ask about this? Everything is fine! What metadata? Everything is fine. It works. And relax. Find me examples of METS with this docStruct! Everything is fine. You need to change every METS file. Please change it. I need your storage projections now. What future plans? This book is 404. I need to buy 4 servers. Why does this happen? Everything is fine. I need your documentation. Have a cup of tea. Unexpected event. Everything is fine. What's for lunch? Please do this. Everything is fine! **OMG IT'S BROKEN!** Everything is fine. Unexpected event. We need to update every METS file. Can it do this? Don't be silly. Training? Priorities have changed. Everything is fine. Everything is fine. Unexpected event. Everything is fine. Why does this happen? Everything is fine. The ALTO isn't defining page headers again. Please change it. Everything is fine. Take a deep breath. You want what! Everything is fine. Can it do this? How's testing on the dev server? Please don't do that again. Please comment on the strategy document. Everything *is* fine! What future plans? That thing happened again. Can you ask about this? It works. Please do this. Budget forecasts have changed. Everything is fine. Why does this happen? New project. Everything is fine. Have a cup of tea. And relax... Zzzzzzzzzzz...

# Quick ~~career~~ round up


1. Degree in history, Masters in Library & Information Studies. (MLIS)
2. Career in information in the UK & New Zealand.
3. Developed information systems/processes for government.
4. National Library of New Zealand, web archiving, systems replacement, policy on born digital material.
5. Wellcome Trust since 2004. Born digital archives & digitisation.
6. 25 years working with information.





# Is a career in information fun...?

1. Introduced to 3 Prime Ministers.
2. Worked with Chief Executives & Cabinet Ministers.
3. Represented my country on UN project.
4. Created some of the first websites for government departments.
5. Independent consultant.
6. Published, travelled, spoken at conferences, reviewed peers work, invited onto external committees.
7. Well, is it fun...?



**Fcuk,  
yes!**

# So, how do I spend my day...?

1. It's not all glamour, metadata & preservation events.
2. Responsible for 2 systems; #Goobi & Preservica  
#preservica (read 'Snagging') & Library storage.
3. Planning. Planning. Planning. Meetings. Meetings.  
Meetings.
4. Proposals. Project management. Deadlines. Budgets.  
Objectives. Reporting. Evaluation.
5. Tiny, tiny, tiny part of my time spent on stuff like data  
recovery, migration, preservation events.



<http://atropapharmacy.co.uk/?product=dispersible-aspirin-tablets-75mg-low-dose-aspirin>

So, six things I  
wish I'd known  
before I  
started...





# 1. Nothing without imagination

1. Data is pretty cool stuff. It's power to change lives can only be revealed if we can imagine the uses the data can be put to.
2. Digital preservation is not an exercise in technology for its own sake. It's a social activity.
3. Digital preservation is about engaging with creators & consumers, with the data & with the future.

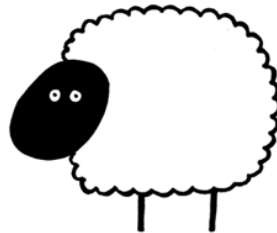
Logic will get you from  
**AtoB**  
imagination will take you  
**everywhere**

- Albert Einstein

<https://dracus.wordpress.com/tag/imagination/>

## 2. No such thing as digital preservation

1. Only point of digital preservation is reuse of data. Nothing else.
2. It's not the ambulance at the bottom of the cliff.
3. It's not about format migration & obsolete kit.
4. It's the little things that we do every day; backup, planning, IT infrastructure maintenance, policy, virus checking, integrity checking, working with creators, working with users, adding value to data.



Don't follow me  
I haven't got a clue

### 3. Lets see how hard we can make this...

1. We say digital preservation is hard. We choose to do this.
2. Much is under our control or within our gift.
3. Unless we can make preservation an easy & sustainable activity it's not preservation.
4. We have to work in the real world.
5. Why choose to make it hard?



## 4. It is a quest for world domination

1. I believe that one day archivists will rule the world.
2. Archivists have all the training to make decisions about selection, appraisal, arrangement, context, description & preservation.
3. Archivists understand preservation, copyright, access, privacy. Stuff that matters.
4. If the amount of personal data that Google holds scares you then the power of archivists to manipulate that data should scare you even more.



Apologies to the  
archivists, you can  
of course trust  
them completely.

## 5. Its not a technical problem...

1. Digital preservation is not a technical problem; it's a social problem.
2. Many technical issues already have solutions, e.g. PLANETS\PLATO addresses many of the issues of migration.
3. The challenge is to move digital data into the future not just intact but in context & retaining its significance as research material that can explain the past.
4. This makes it a social problem.

## 6. That I'd start so many arguments!

1. The power of influence is shifting & becoming democratized.
2. Anyone can have 15mb of fame; and that's a good thing!
3. Have an opinion. Have a thought. Have an idea. Share it. Tweet it. Blog it.
4. Not because you want to build a career, but because you give a damn.

# I wish I'd known that...

1. I'd be making opportunities to be creative & to use my imagination, but also to be challenged.
2. I'd be building something that has value & worth.
3. I'd be engaging with some astonishing people & even more astonishing material.
4. I'd be involved in the making of history.



I need you to  
know this  
before you  
start.

Firefox

BBC News - Conservatives purge old spe... +

www.bbc.co.uk/news/uk-politics-24924185

Goobi - production Goobi - 2.0 SDB - production DDS Dashboard ITM FILEX Test SDB ITM II

Sign in News Sport Weather iPlayer TV Radio More... Search

# NEWS POLITICS

Home World UK England N. Ireland Scotland Wales Business Politics Health Education Sci/Environment Technology Entertainment & Arts

13 November 2013 Last updated at 15:59

## Conservatives purge old speeches from online archives

The Conservative Party has deleted speeches and press releases published on its website between 2000 and the 2010 general election.

The archive has also been hidden from search engines.

The move was spotted by Computer Weekly, a trade publication, which also said some records had been removed from the Internet Archive, which aims to make a permanent record of web content.

The Conservatives said the move keeps their revamped website up-to-date.

Computer Weekly said the effect of the changes was "as alarming as sending Men in Black to strip history books from a public library and burn them in the car park".

"Prime minister David Cameron and Chancellor George Osborne campaigned on a promise to democratise information held by those in power, so people could hold them to account," it added.

### 'Broken promises'

Instead, they have overseen a change which has left Conservative speeches in a "secretive corner of the internet like those that shelter the military" and "gangsters", it concluded.

Archived pages on the Conservative Party website providing a record of evolving policy announcements were deleted "sometime after 5 October", Computer Weekly said.

In addition, a text file was added to the website, called "robots.txt". This is a standard way of telling search engines which parts of a website they should not try to index.

But the website did not have the effect of preventing search engines from



"By making more information available to more people, you're giving them more power," the PM said in 2006

### Top stories

Hospitals to publish staffing levels

Toronto mayor stripped of powers

'Selfie' named as word of the year

Independence to 'transform' economy

Deadly cyclone strikes Sardinia

### Features

**Secrets of the Tardis**  
Step inside and explore the Doctor Who time ship

**Death toll**  
Is cycling getting more dangerous or becoming safer?

**Playboy defector**  
The city with good memories of Lee Harvey Oswald

**Second chance**  
The man who has a recycled pacemaker

### Most popular

Shared Read Video/Audio

US zoo shock as lion kills lioness	1
'Last' tank to be used in Pitt film	2
'Selfie' named as word of the year	3
Zimmerman 'pointed shotgun at woman'	4

On the UK Web Archive held by the BL...

"It's not a public  
archive. It's an archive  
for preservation and for  
research,"

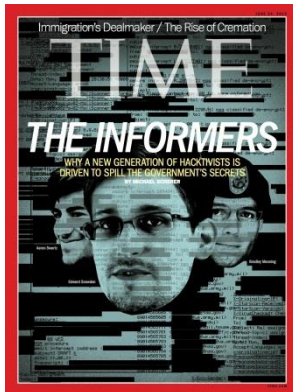
Angela Mills Wade, who helped formulate the 2013 regulations as executive director of the European Publishers Council, said it was not necessarily in the public interest that the internet archive should be made generally available.

<http://www.computerweekly.com/news/2240210795/UK-prepares-to-launch-internet-archive-without-internet-access>

Don't get mad.  
Just do your job.

Doing your job without  
fear or favour is what  
the future will thank  
you for.

# One of you may be the next Edward Snowden.



Would that be such a  
bad thing?

Questions now, questions later...?

Dave Thompson

Digital Curator  
Wellcome Library

[d.thompson@wellcome.ac.uk](mailto:d.thompson@wellcome.ac.uk) @D\_N\_T