

# The Sudamih Project: Findings and Conclusions

Monday 28 March 2011

James A J Wilson

JISC

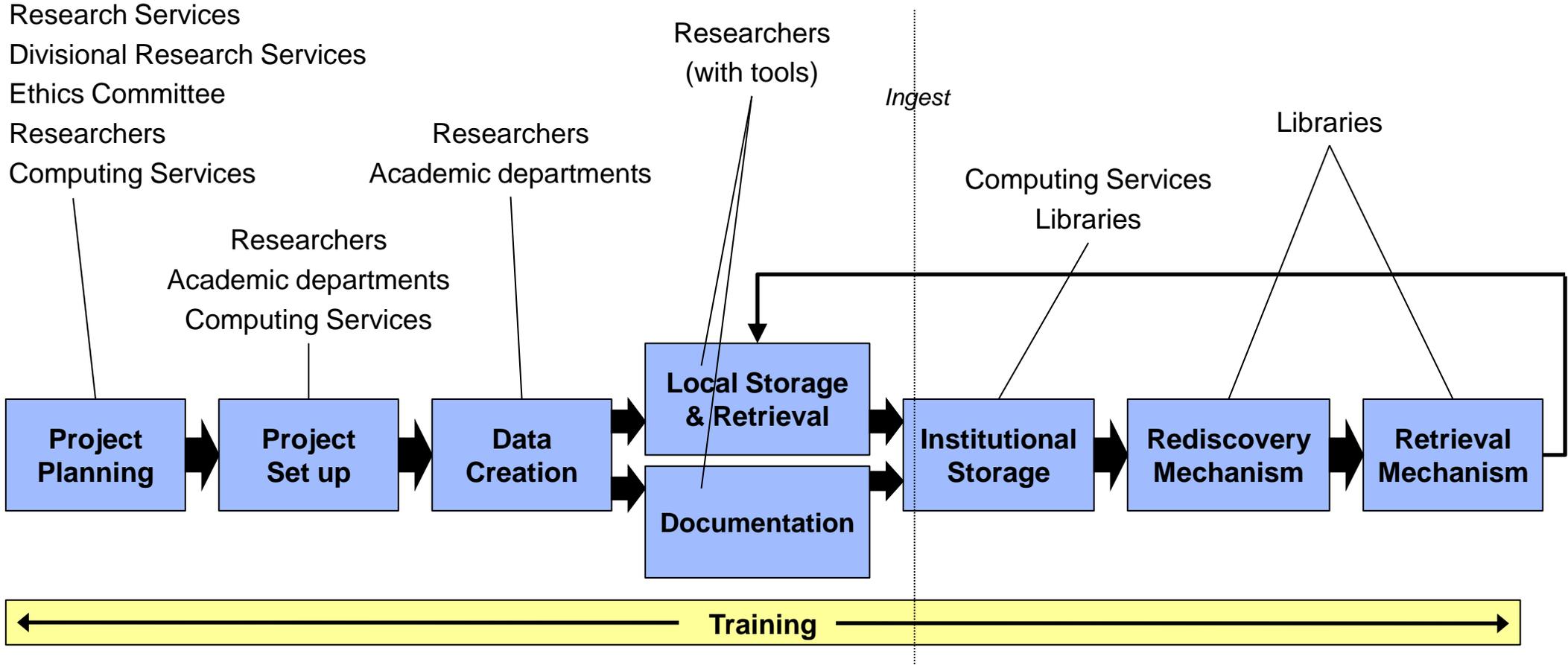


UNIVERSITY OF  
OXFORD

# Sudamih

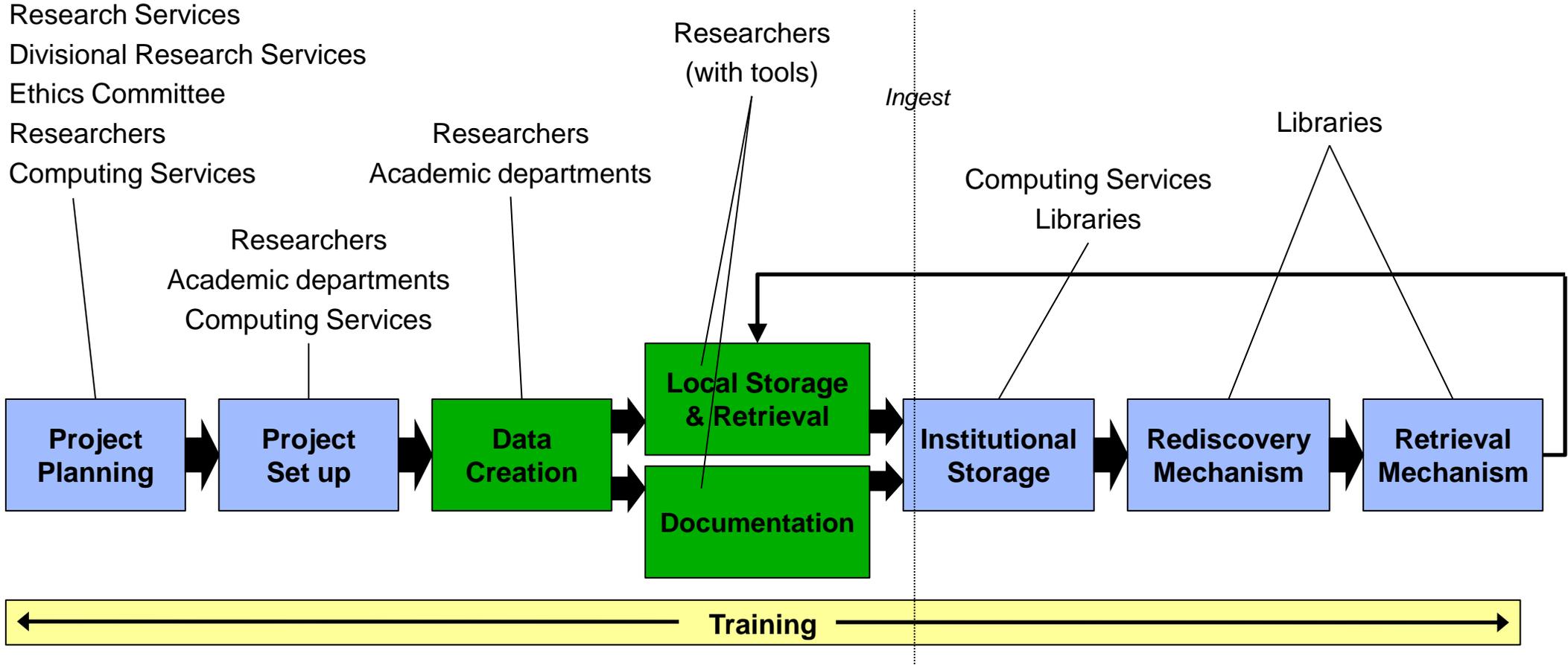
- Supporting Data Management Infrastructure for the Humanities
- Part of a programme of data management infrastructure development at the University of Oxford
  - Scoping Digital Repository Services for Research Data Management (internally funded)
  - EIDCSR (Embedding Institutional Data Curation Services in Research) (JISC funded)
- Intention to extend disciplinary coverage and create integrated research data management processes

# Phases of Research Data Management



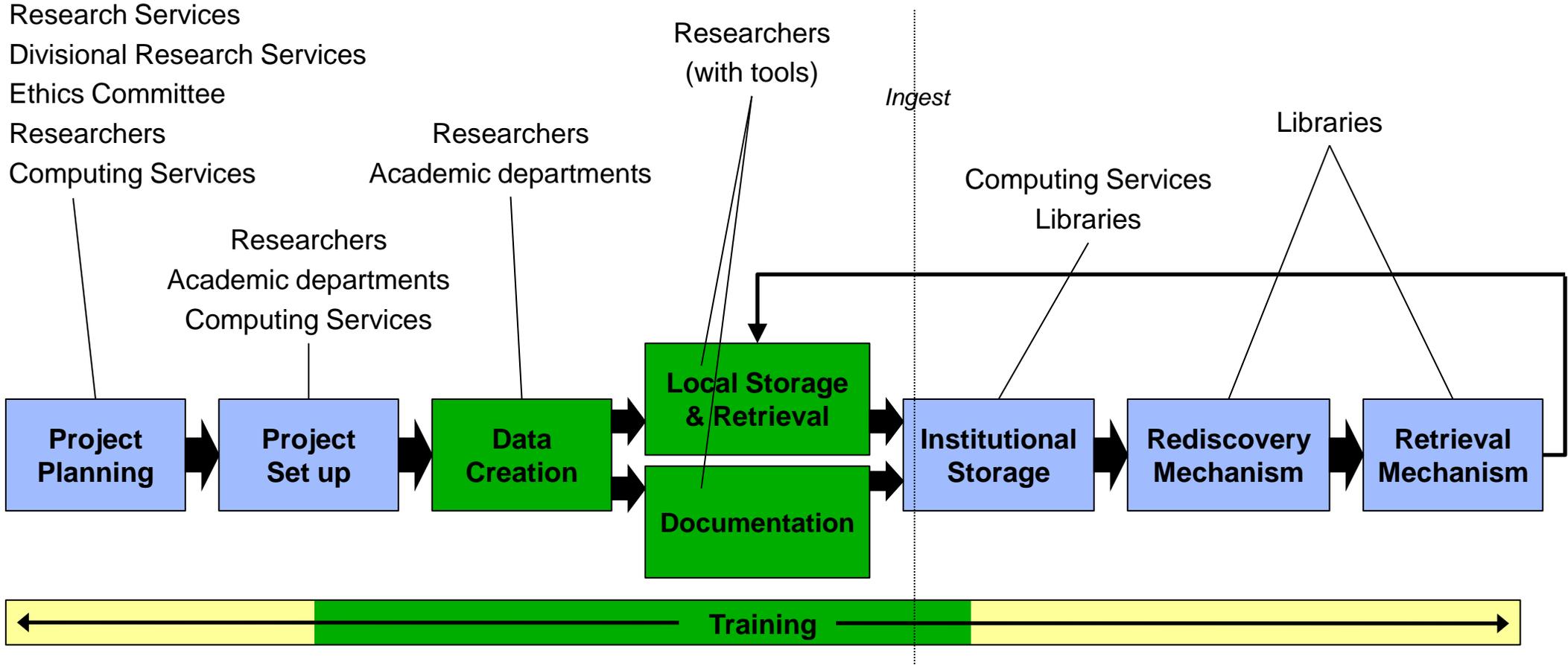
Research Services; Divisional Research Services; academic divisions; academic faculties; Computing Services; Libraries.

# Database as a Service



Research Services; Divisional Research Services; academic divisions; academic faculties; Computing Services; Libraries.

# Data Management Training



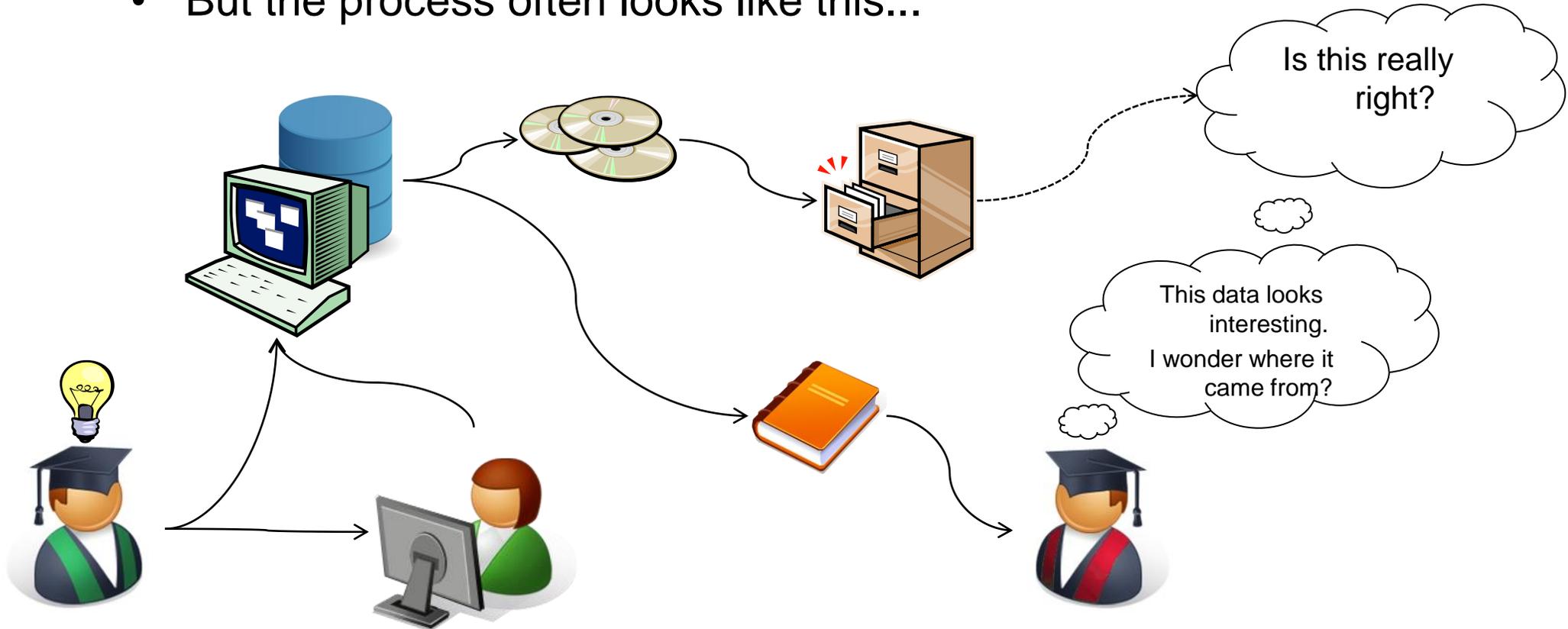
Research Services; Divisional Research Services; academic divisions; academic faculties; Computing Services; Libraries.

# Information/data Management Practices

- Little consideration given to data management
- Diverse and idiosyncratic data management practices
  - Lone scholar tradition contributes to this
- Low awareness of central services
- Little or no prior training

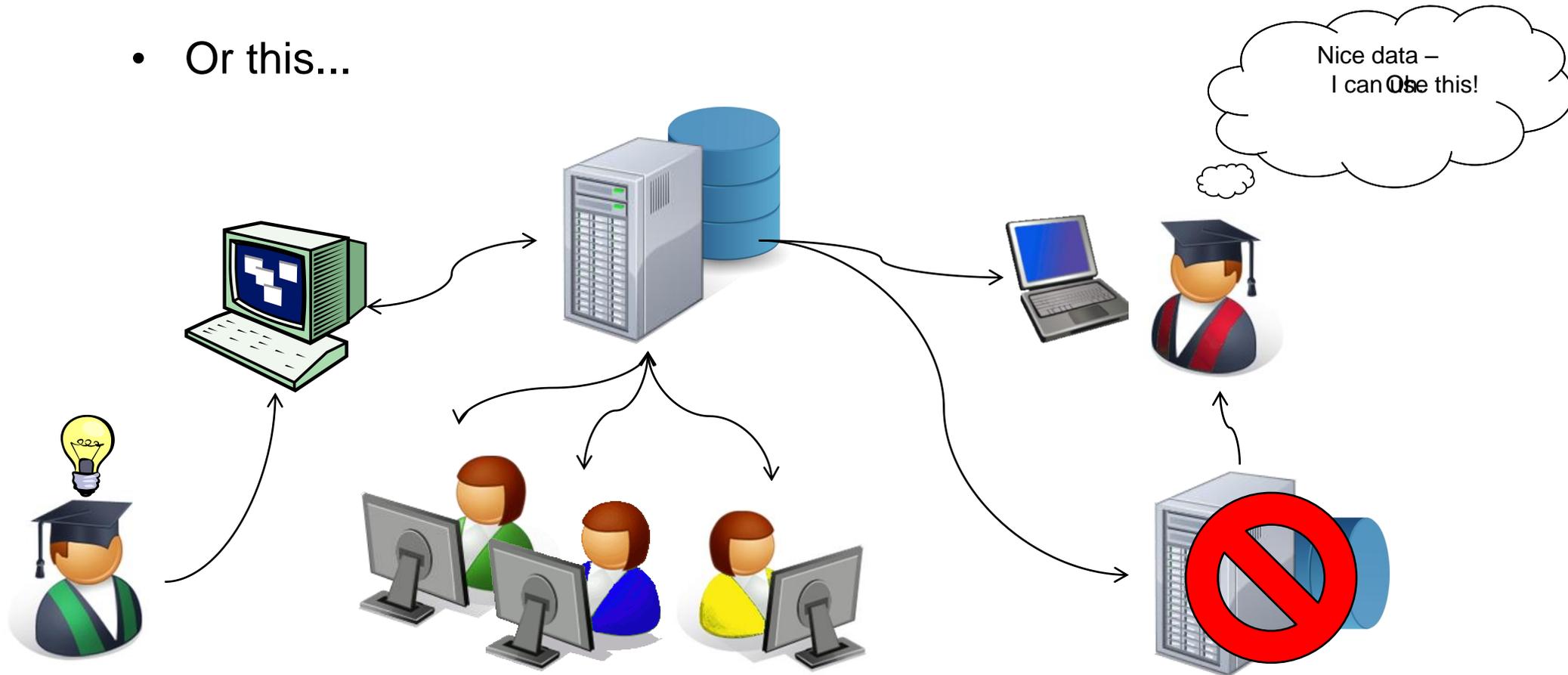
# Databases in the humanities – current practices

- Database practices vary greatly in the humanities
- But the process often looks like this...



# Databases in the humanities – current practices

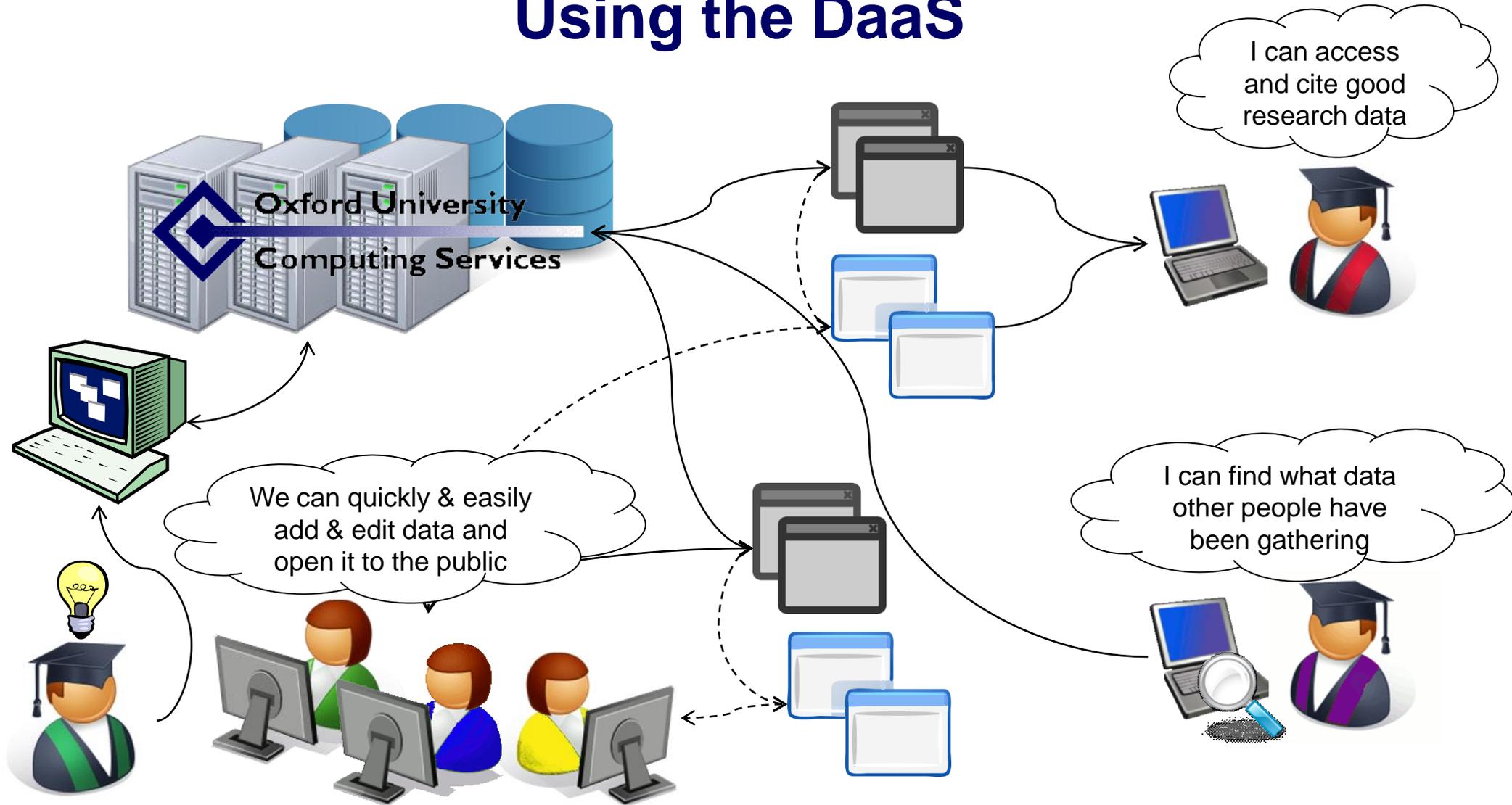
- Database practices vary greatly in the humanities
- Or this...



# Problems identified

- Lack of technological awareness
- Poor backing-up practices
- Collaboration difficult
- Difficult to re-discover and re-use data
- Risk of technical obsolescence
- Money and long-term sustainability
- Technical expertise required
- What happens to data when its creators move?

# Using the DaaS

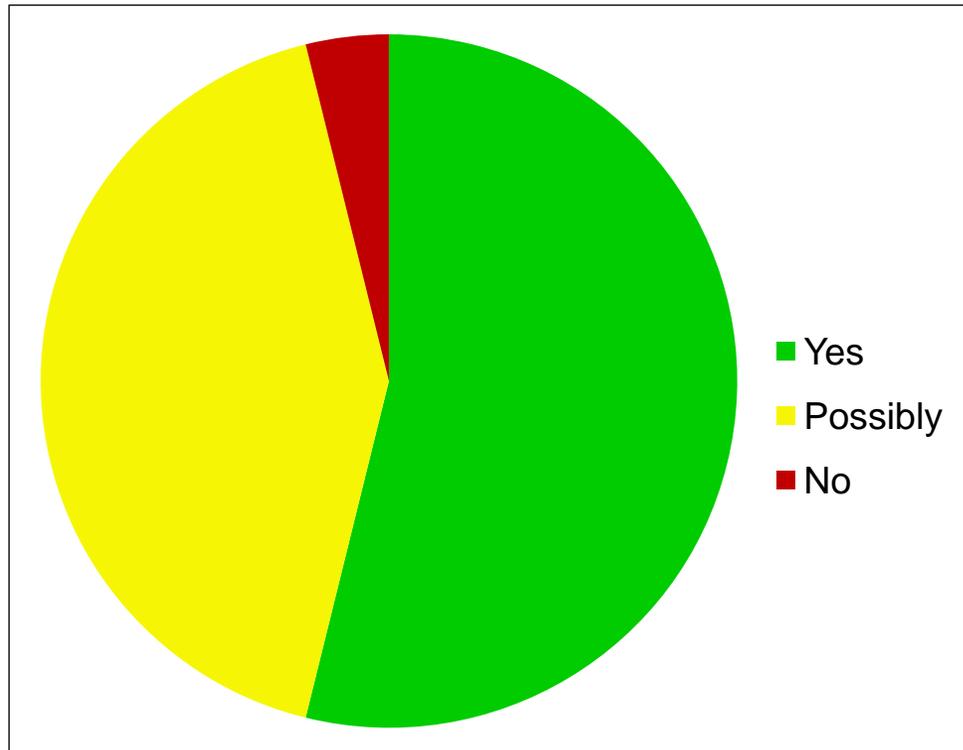


# Future possibilities

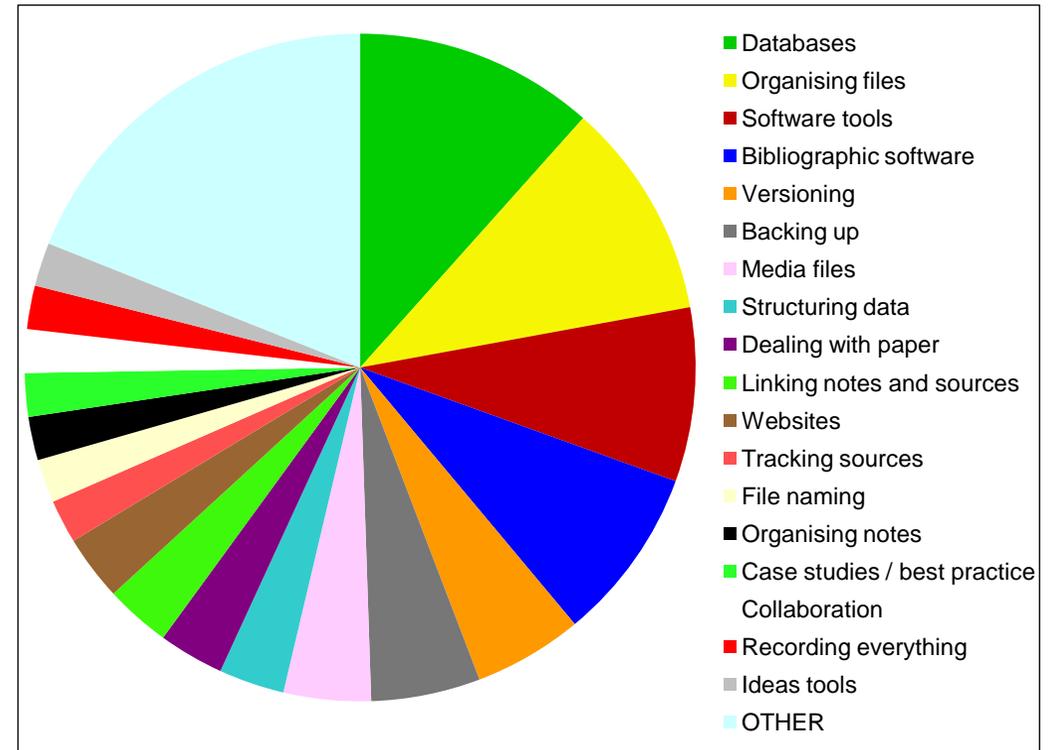
- Full productionization and scalability
- Registry service for descriptive metadata
- Extended functionality beyond the humanities
- Incorporation of data storage models other than relational databases
  - e.g. XML-based, document-based, and CouchDB
- Deployment as a cloud-based software service
  - Involves advance monitoring and management tools
  - Capable of running on other institutional virtual infrastructures
- End-user documentation and support

# Demand for training

Is there a need for data management training?



What should this training cover?



# Approach to Training

1. **Introduction to data management.** What is data management? What tools and services does Oxford provide to help do it?
2. **Tools to help manage research data.** Considering real-life research challenges and problems faced – which tools/methods are best for solving them? Spreadsheets; databases; xml; bibliographic software; other ‘research’ software; how to structure and query data.
3. **Organising and linking research information for later retrieval.** Including information on organising paper-based as well as all forms of electronic sources (notes, journal articles, books, references, images, numerical data, multimedia). Includes: versioning; file & folder structures; classification; linking sources and themes; making things searchable.
4. **Technical aspects of funding bids.** How to plan and write about the technical aspects of research bids; examples of successful bids, technical advisory, IPR, long-term data curation.
5. **Database design for humanities research data.** Associating people, places, things, and events; distinguishing entities with shared names; GIS data, uncertain dates, incomplete data, non-Roman alphabets, recording sources.

# Training/support Materials Developed

## RESEARCH DATA MANAGEMENT UAS

Enter search term Search

[The site](#)
[University of Oxford](#)
[People](#)

UAS Home > Research Data Management >

- Why manage your data?
- Data Management Planning
- Data Backup and Security
- Data Sharing and Archive
- Training, Advice & Support

### Research Data

Good practice in data management, integrity, or the responsible use of research data.

The following diagram provides an overview of research data management, both within the University and across the globe.

### Managing Research Information

Recent pages: [Managing research information](#) > [Keep your research information safe](#)

Whatever the nature of your research project, you will generate information of various sorts. It's important to manage this information in a way that makes it easy to find, and an important part of efficient research methods.

- Managing references, citations and bibliographies
- Organising research material
- Synchronising files
- Keeping your data safe
- Managing structured data
- Emailing
  - With a good understanding of how to use email
  - Nexus - for Oxford email

About the Research Data Management website  
[Contact us](#) / [Report a problem with the website](#)

## Research Skills Toolkit

### Choosing Bibliographic Software

Bibliographic software allows you to create and manage bibliographic records for the books, journals, articles, and other research. This can be used to add citations and references (in a wide variety of styles) to your own writing. You can also use it to manage records in your data management software.

Many different packages are available. Although the details vary, and selecting a package often depends on the features that are important to you, the following are some of the most popular tools: *EndNote*, *RefWorks*, *Zotero*, and *Mendeley*.

#### Compatibility

*EndNote* is available for Windows and Mac. It is also compatible with Linux. All three can be used with *Microsoft Word* and *Open Office.org Writer* documents. *RefWorks* is a web-based service, and therefore works with any operating system. However, *Write-N-Cite*, the *RefWorks* plug-in for *Microsoft Word*, the citation process is a little more complex than the other word processors.

#### Free or Paid For?

*EndNote* is a commercial program. If you purchase the full desktop version, it currently costs about £80. However, *EndNote Web* is free, but has limited features, and is intended for use with *EndNote* desktop software.

*RefWorks* is available free of charge to Oxford University researchers via an institutional subscription.

*Zotero* is free and open source. Additional online services are available for a fee.

#### Online or Offline?

*EndNote* is a desktop application which you run on your computer. It can be used for offline working with local databases, or for online working with online services. *EndNote Web* is an online service accessed by logging in to the *EndNote Web* website. References are stored on a *RefWorks* server, and you access them from any computer with a Web connection. For *EndNote*, you download a copy of your references and use them locally.

## What is research data management?

### Research Information Management: Tools for the Humanities

Download and install *Tablets*

View files and folders using *Tablets*

Create tablets (or bubbles) and add files to them

Search for files using *Tablets*

All the files for these exercises have been provided for you as a network drive. The drive is called 'R1', and your area of the drive is called 'IT Learning Programme'. Some software packages will refer to it as 'ras', some as 'file server'.

Among the files provided is one called 'File Tool website links.pdf'. This provides clickable links to the home pages of the tools described in this course book, to save you from having to type each URL individually.

Task 1: Download and install <i>Tablets</i>	Step 1: Start the computer if necessary.
	Step 2: Click the Start button on the Task Bar at the bottom of the screen. Hold the mouse pointer over All Programs, and select your choice of where to search for programs: My Recent Places, All Programs, or All Programs   Accessories. Click on the link for <i>Tablets</i> .
	Step 3: Open the File Tool website links.pdf (you can do this in Windows Explorer, which can be found by clicking the Start button, then going to All Programs   Accessories). Click on the link for <i>Tablets</i> . Alternatively, type <a href="http://tablets.ox.ac.uk/Software/Download/Installation.html">http://tablets.ox.ac.uk/Software/Download/Installation.html</a> in your browser's address bar, and press Enter.
	Step 4: Download and install the software. Follow the on-screen instructions to download and install the program.
Task 2: Launch <i>Tablets</i>	Step 1: Click the Start button on the Task Bar at the bottom of the screen.

IT Learning Programme

UNIVERSITY OF OXFORD

### Research Information Management: Tools for the Humanities

specific combination of tags, or even to view the files that don't have a particular tag.

If you create a new file or modify an existing one while *Tablets* is running, the program will give you the opportunity to add tags when you save it.

The basic version of *Tablets* can be downloaded free of charge; previous versions which offer more features are available on payment of a fee. See <http://tablets.ox.ac.uk/> for more information.

Exercise 1: Using *Tablets* to organise files

- Download and install *Tablets*
- View files and folders using *Tablets*
- Create tablets (or bubbles) and add files to them
- Search for files using *Tablets*

All the files for these exercises have been provided for you as a network drive. The drive is called 'R1', and your area of the drive is called 'IT Learning Programme'. Some software packages will refer to it as 'ras', some as 'file server'.

Among the files provided is one called 'File Tool website links.pdf'. This provides clickable links to the home pages of the tools described in this course book, to save you from having to type each URL individually.

Task 1: Download and install *Tablets*

Step 1: Start the computer if necessary.

Step 2: Click the Start button on the Task Bar at the bottom of the screen. Hold the mouse pointer over All Programs, and select your choice of where to search for programs: My Recent Places, All Programs, or All Programs | Accessories. Click on the link for *Tablets*.

Step 3: Open the File Tool website links.pdf (you can do this in Windows Explorer, which can be found by clicking the Start button, then going to All Programs | Accessories). Click on the link for *Tablets*. Alternatively, type <http://tablets.ox.ac.uk/Software/Download/Installation.html> in your browser's address bar, and press Enter.

Step 4: Download and install the software. Follow the on-screen instructions to download and install the program.

Task 2: Launch *Tablets*

Step 1: Click the Start button on the Task Bar at the bottom of the screen.

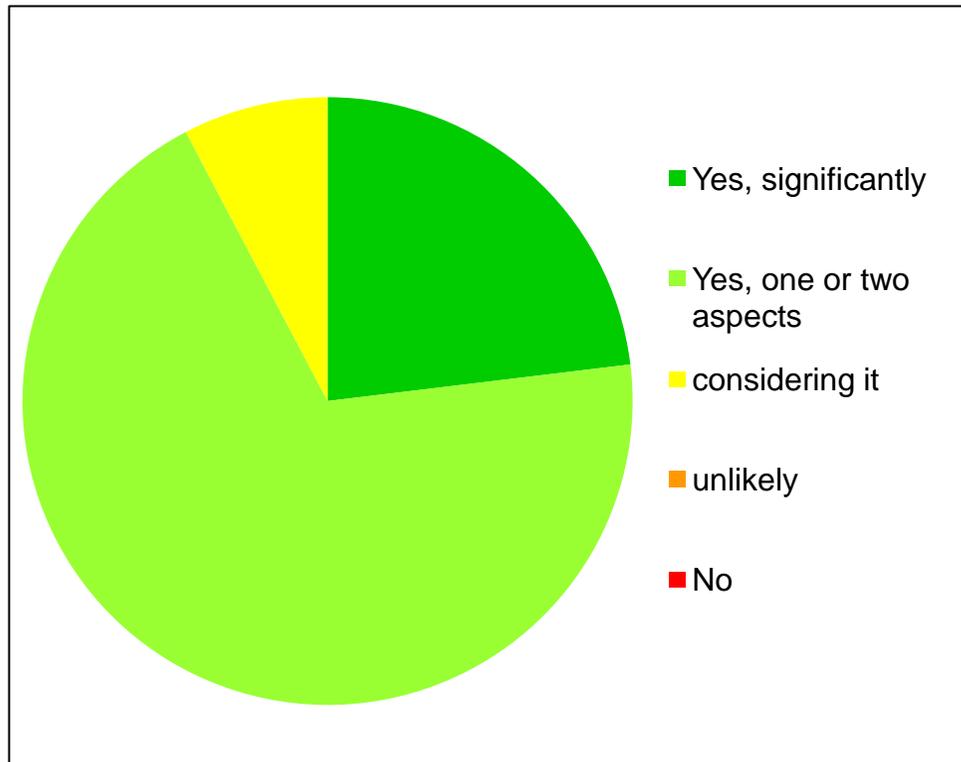
IT Learning Programme offers a wide range of courses – for both beginners and more experienced users. There are courses on specific software packages and on more general skills such as design, working with digital images, and so on, including some on data management. [File Tools](#) offers links to other University training providers.

For more information and advice, the University of Oxford [Research Data Management website](#) provides guidance and further information about the services available. A [leaflet](#) covering the key issues can be downloaded from [Research Data Management Project's website](#).

THIS WEBSITE IS PROVIDED BY THE [UNIVERSITY OF OXFORD LIBRARY SERVICES](#) AND IS MADE AVAILABLE UNDER THE TERMS OF THE [UNIVERSITY OF OXFORD LIBRARY SERVICES' OPEN ACCESS LICENSE](#) (<http://www.library.ox.ac.uk/About/About-us>)

# Response to Courses

Have you or will you change any aspects of your own information management practices as a result of the course?



- I will look at IT tools from a different perspective (how they will move my project forward, rather than how I fit my project to the IT tools)
- After the course I started to work as a 'horizontal organizer' and I feel it works much better for me. Also now I plan ahead how to get these materials organized and traceable, rather than working it out afterwards as I did before.
- I'm getting more and more organized and happy now! It changed not only my ways of working, but also my mood and my attitude

# Costs & Benefits of Training

1. Time saved by researchers by locating and retrieving relevant research notes and information more rapidly
2. Improved quality of research by locating better, more relevant research information than would otherwise be the case
3. Improved quality of research by linking materials in such a way as to highlight connections and trigger new ideas
4. Improved comprehensibility of research information and data after long time periods
5. Better awareness and use of software tools to assist research management
6. Better awareness and uptake of central infrastructure services intended to help researchers, including technical help and assistance with funding bids
7. Reduced risk of data loss
8. Improved version control

Activity	Resource	Costs per term (to run both courses once)
Update course content	25 hours at £13.77	£344.25
OUCS Training Room hire & set-up	25-person capacity room for 6 hours. Not charged to ITLP-led courses.	£0 (would otherwise cost c. £180)
Printing of course booklets	25 booklets, covered by course price.	£0 (would otherwise cost c. £150)
Cost of teacher (preparation)	6 hours at £13.77	£83.62
Cost of teacher (delivery)	6 hours at £27.54	£165.24
Cost of assistant (preparation & support)	12 hours at £13.77	£165.24
<b>TOTAL</b>		<b>£758.35</b> <b>(£2,275 annually)</b>

# Costs & Benefits of Training

1. Time saved by researchers by locating and retrieving relevant research notes and information more rapidly
2. Improved quality of research by locating better, more relevant research information than would otherwise be the case
3. Improved quality of research by linking materials in such a way as to highlight connections and trigger new ideas
4. Improved comprehensibility of research information and data after long time periods
5. Better awareness and use of software tools to assist research management
6. Better awareness and uptake of central infrastructure services intended to help researchers, including technical help and assistance with funding bids
7. Reduced risk of data loss
8. Improved version control

Activity	Resource	Costs per term (to run both courses once)
Update course content	25 hours at £13.77	£344.25
OUCS Training Room hire & set-up	25-person capacity room for 6 hours. Not charged to ITLP-led courses.	£0 (would otherwise cost c. £180)
Printing of course booklets	25 booklets, covered by course price.	£0 (would otherwise cost c. £150)
Cost of teacher (preparation)	6 hours at £13.77	£83.62
Cost of teacher (delivery)	6 hours at £27.54	£165.24
Cost of assistant (preparation & support)	12 hours at £13.77	£165.24
<b>TOTAL</b>		<b>£758.35</b> <b>(£2,275 annually)</b>

**25%**

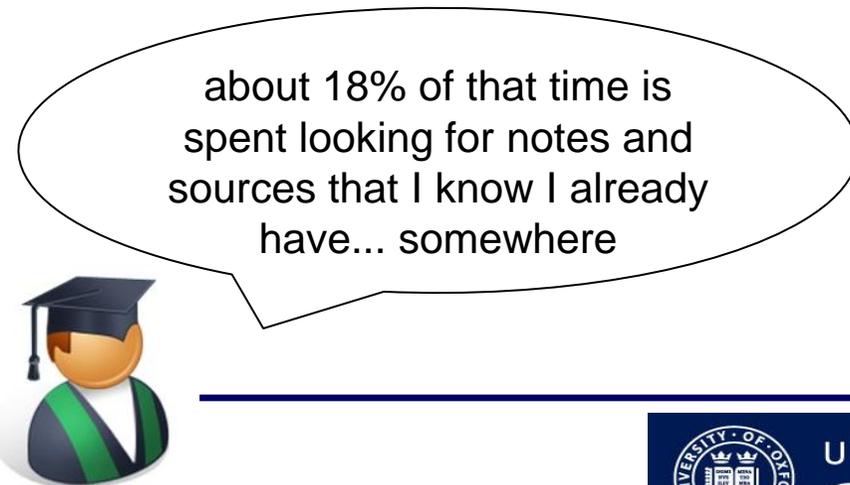
I spend about 25% of my time writing up research publications



# Costs & Benefits of Training

1. Time saved by researchers by locating and retrieving relevant research notes and information more rapidly
2. Improved quality of research by locating better, more relevant research information than would otherwise be the case
3. Improved quality of research by linking materials in such a way as to highlight connections and trigger new ideas
4. Improved comprehensibility of research information and data after long time periods
5. Better awareness and use of software tools to assist research management
6. Better awareness and uptake of central infrastructure services intended to help researchers, including technical help and assistance with funding bids
7. Reduced risk of data loss
8. Improved version control

Activity	Resource	Costs per term (to run both courses once)
Update course content	25 hours at £13.77	£344.25
OUCS Training Room hire & set-up	25-person capacity room for 6 hours. Not charged to ITLP-led courses.	£0 (would otherwise cost c. £180)
Printing of course booklets	25 booklets, covered by course price.	£0 (would otherwise cost c. £150)
Cost of teacher (preparation)	6 hours at £13.77	£83.62
Cost of teacher (delivery)	6 hours at £27.54	£165.24
Cost of assistant (preparation & support)	12 hours at £13.77	£165.24
<b>TOTAL</b>		<b>£758.35</b> <b>(£2,275 annually)</b>

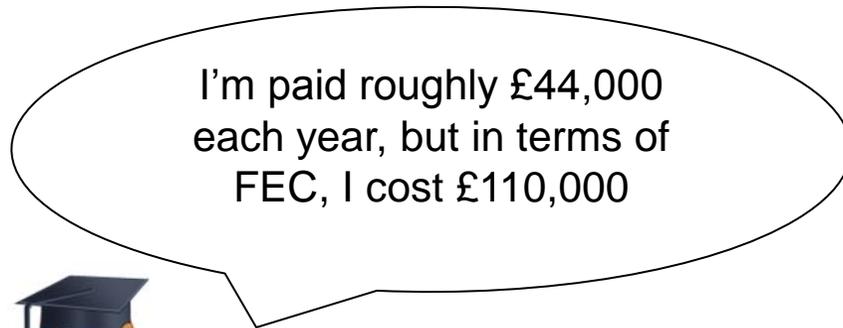


$$25\% \times 18\% = 4.5\%$$

# Costs & Benefits of Training

1. Time saved by researchers by locating and retrieving relevant research notes and information more rapidly
2. Improved quality of research by locating better, more relevant research information than would otherwise be the case
3. Improved quality of research by linking materials in such a way as to highlight connections and trigger new ideas
4. Improved comprehensibility of research information and data after long time periods
5. Better awareness and use of software tools to assist research management
6. Better awareness and uptake of central infrastructure services intended to help researchers, including technical help and assistance with funding bids
7. Reduced risk of data loss
8. Improved version control

Activity	Resource	Costs per term (to run both courses once)
Update course content	25 hours at £13.77	£344.25
OUCS Training Room hire & set-up	25-person capacity room for 6 hours. Not charged to ITLP-led courses.	£0 (would otherwise cost c. £180)
Printing of course booklets	25 booklets, covered by course price.	£0 (would otherwise cost c. £150)
Cost of teacher (preparation)	6 hours at £13.77	£83.62
Cost of teacher (delivery)	6 hours at £27.54	£165.24
Cost of assistant (preparation & support)	12 hours at £13.77	£165.24
<b>TOTAL</b>		<b>£758.35</b> <b>(£2,275 annually)</b>



I'm paid roughly £44,000 each year, but in terms of FEC, I cost £110,000

**25% X 18%**  
**= 4.5%**  
**Of £110,000**  
**= £4,950 per year**

# Costs & Benefits of Training

1. Time saved by researchers by locating and retrieving relevant research notes and information more rapidly
2. Improved quality of research by locating better, more relevant research information than would otherwise be the case
3. Improved quality of research by linking materials in such a way as to highlight connections and trigger new ideas
4. Improved comprehensibility of research information and data after long time periods
5. Better awareness and use of software tools to assist research management
6. Better awareness and uptake of central infrastructure services intended to help researchers, including technical help and assistance with funding bids
7. Reduced risk of data loss
8. Improved version control

Activity	Resource	Costs per term (to run both courses once)
Update course content	25 hours at £13.77	£344.25
OUCS Training Room hire & set-up	25-person capacity room for 6 hours. Not charged to ITLP-led courses.	£0 (would otherwise cost c. £180)
Printing of course booklets	25 booklets, covered by course price.	£0 (would otherwise cost c. £150)
Cost of teacher (preparation)	6 hours at £13.77	£83.62
Cost of teacher (delivery)	6 hours at £27.54	£165.24
Cost of assistant (preparation & support)	12 hours at £13.77	£165.24
<b>TOTAL</b>		<b>£758.35</b> <b>(£2,275 annually)</b>

If I attend the two training courses and, as a result, I only spend 16% of my time hunting notes & sources

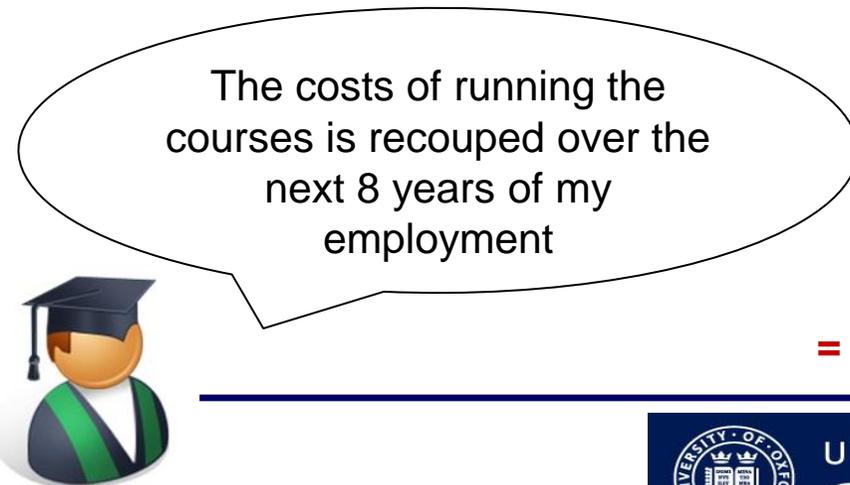


**25% X 18%**  
**= 4.5%**  
**Of £110,000**  
**= £4,950 per year**  
**2% of £4,950**  
**= £99**

# Costs & Benefits of Training

1. Time saved by researchers by locating and retrieving relevant research notes and information more rapidly
2. Improved quality of research by locating better, more relevant research information than would otherwise be the case
3. Improved quality of research by linking materials in such a way as to highlight connections and trigger new ideas
4. Improved comprehensibility of research information and data after long time periods
5. Better awareness and use of software tools to assist research management
6. Better awareness and uptake of central infrastructure services intended to help researchers, including technical help and assistance with funding bids
7. Reduced risk of data loss
8. Improved version control

Activity	Resource	Costs per term (to run both courses once)
Update course content	25 hours at £13.77	£344.25
OUCS Training Room hire & set-up	25-person capacity room for 6 hours. Not charged to ITLP-led courses.	£0 (would otherwise cost c. £180)
Printing of course booklets	25 booklets, covered by course price.	£0 (would otherwise cost c. £150)
Cost of teacher (preparation)	6 hours at £13.77	£83.62
Cost of teacher (delivery)	6 hours at £27.54	£165.24
Cost of assistant (preparation & support)	12 hours at £13.77	£165.24
<b>TOTAL</b>		<b>£758.35</b> <b>(£2,275 annually)</b>

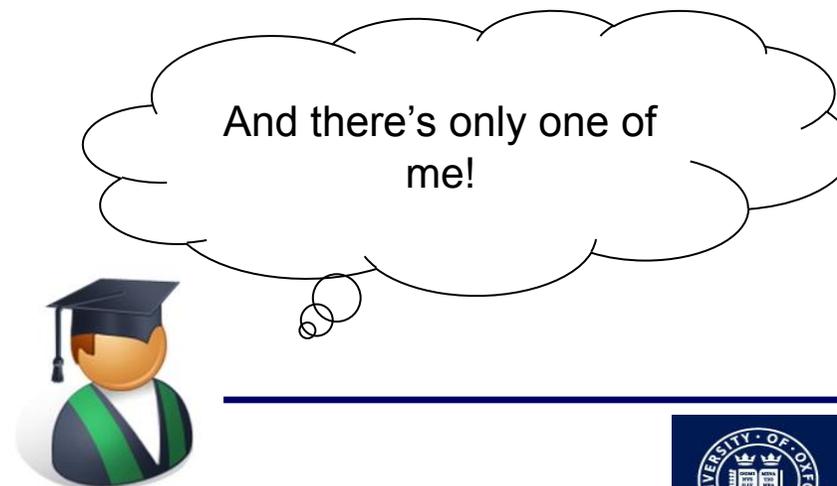


**25% X 18%**  
**= 4.5%**  
**Of £110,000**  
**= £4,950 per year**  
**2% of £4,950**  
**= £99 x 8 years = £792**

# Costs & Benefits of Training

1. Time saved by researchers by locating and retrieving relevant research notes and information more rapidly
2. Improved quality of research by locating better, more relevant research information than would otherwise be the case
3. Improved quality of research by linking materials in such a way as to highlight connections and trigger new ideas
4. Improved comprehensibility of research information and data after long time periods
5. Better awareness and use of software tools to assist research management
6. Better awareness and uptake of central infrastructure services intended to help researchers, including technical help and assistance with funding bids
7. Reduced risk of data loss
8. Improved version control

Activity	Resource	Costs per term (to run both courses once)
Update course content	25 hours at £13.77	£344.25
OUCS Training Room hire & set-up	25-person capacity room for 6 hours. Not charged to ITLP-led courses.	£0 (would otherwise cost c. £180)
Printing of course booklets	25 booklets, covered by course price.	£0 (would otherwise cost c. £150)
Cost of teacher (preparation)	6 hours at £13.77	£83.62
Cost of teacher (delivery)	6 hours at £27.54	£165.24
Cost of assistant (preparation & support)	12 hours at £13.77	£165.24
<b>TOTAL</b>		<b>£758.35</b> <b>(£2,275 annually)</b>



$$\begin{aligned}
 &25\% \times 18\% \\
 &= 4.5\% \\
 &\text{Of } \pounds 110,000 \\
 &= \pounds 4,950 \text{ per year} \\
 &2\% \text{ of } \pounds 4,950 \\
 &= \pounds 99 \times 8 \text{ years} = \pounds 792
 \end{aligned}$$

# Lessons learnt

- Mind your language
- Communication and engagement is important. Need buy-in at multiple levels
- Researchers are key, but they probably don't think all that much about re-usability, long-term curation, etc.
- Understand the rhythms of the academic year
- Use existing infrastructure where possible
- You may need to compromise when embedding data management training into existing training

Thanks!