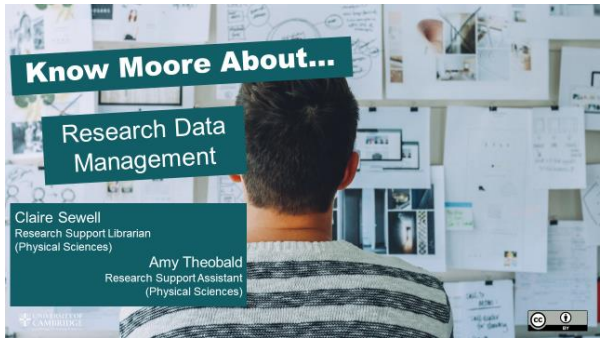
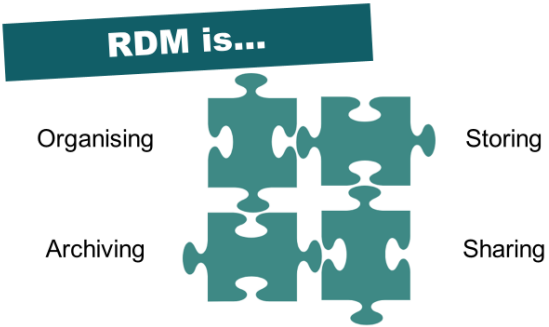



Know Moore About: RDM


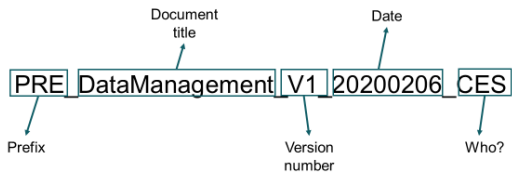
	<p>Script:</p> <ul style="list-style-type: none">• Welcome to the session and introductions.• Housekeeping.• Slides/handouts will be distributed electronically.
<p>Presenter notes:</p>	<p>Script:</p> <ul style="list-style-type: none">• Topics to cover in this session.• Look at what data actually are, how they can be organised and stored.• Focus on the vital role of open data in open research – especially when it comes to the requirements to share data.• Very brief tour of the basics of data management – can point you towards more information if needed.
<p>Topics</p> <ul style="list-style-type: none">▪ Introduction to data▪ Organising data▪ Storing data▪ Archiving data▪ Data sharing	
<p>Presenter notes:</p>	

<div data-bbox="204 226 807 562"> <h3>What is data?</h3> <ul style="list-style-type: none"> Documents Lab books Videos Field notes Experiment results Images Protocols Survey Manuscripts Interview transcripts Physical samples <p>Anything used and produced during research</p> </div>	<p>Script:</p> <ul style="list-style-type: none"> Data can mean many things – often very discipline dependent. Term covers any of the terms highlighted on the screen and more. Some people struggle as seeing what they work with as data. This is more of an issue in HASS. Data = any information used or produced during the research process. 								
<p>Presenter notes:</p> <p>Interaction – ask the audience what <i>data</i> means to them.</p> <p>Animated slide - click to see data examples.</p> <p>Animated slide - click to see data statement.</p>									
<div data-bbox="204 969 807 1305"> <h3>FAIR Data</h3> <p>Helps data to be discovered by both humans and machines</p> <table> <tr> <td>F</td> <td>Findable</td> </tr> <tr> <td>A</td> <td>Accessible</td> </tr> <tr> <td>I</td> <td>Interoperable</td> </tr> <tr> <td>R</td> <td>Reusable</td> </tr> </table> </div>	F	Findable	A	Accessible	I	Interoperable	R	Reusable	<p>Script:</p> <ul style="list-style-type: none"> All research data should follow the FAIR principles – helps data to be found by both researchers and machines = makes it more discoverable. Originally published in 2016 to act as a guideline for those wanting to make their data reusable. Findable – making the data discoverable e.g. adding metadata and assigning a DOI. Accessible – ensuring people can actually use it when found e.g. using recognised systems and open protocols. Interoperable – making sure that data will work across systems e.g. using open source software. Reusable – ensuring data can be re-used and built upon by others e.g. adding an appropriate licence.
F	Findable								
A	Accessible								
I	Interoperable								
R	Reusable								
<p>Presenter notes:</p>									

	<p>Script:</p> <ul style="list-style-type: none"> • Good data management is like building a jigsaw – you need all the right pieces to put together to make a good strategy. • There are some different ways to put a data strategy together and you may spend more time on some areas than others but this will form the basis of any good data plan.
<p>Presenter notes:</p>	
	<p>Script:</p> <ul style="list-style-type: none"> • Begin by looking at how data can be organised. • Fairly basic and likely to be things you do anyway but you will need to think about them and record them in any formal plan. • Useful tips for organising information more generally.
<p>Presenter notes:</p>	



<div data-bbox="204 226 805 564" data-label="Image"> </div>	<p>Script:</p> <ul style="list-style-type: none"> • Almost all researchers work with some type of digital file. • In theory these are easy to find as you can just search for the name of a file but this doesn't always work. • Seems like Example B is the better of the two but each has pros and cons. • Example A – useful for finding recently used files quickly or saving things on the run BUT is messy and will be of limited use in time or if new people come into the project. • Example B – much more structured and easy to follow for anyone coming into the project but could become complex with many levels. • It's about having a system that works for you and your team that is organised in some way. •
<p>Presenter notes:</p> <p>Interaction – ask the audience for the pros/cons of each model and why they think this.</p>	
<div data-bbox="204 1061 805 1397" data-label="Image"> </div>	<p>Script:</p> <ul style="list-style-type: none"> • Many people also deal with physical samples and principles are broadly the same. • Create maps – can be simple as a spreadsheet. Keep a note of where things were found/what they were used for. Add relevant notes. • Reference – create references for all samples and keep these up to date. Can be useful if you need to restock or when keeping variable the same. • Paper notes – can be converted to digital notes by apps or scanning. Can help to keep things safe and easier to share later on.
<p>Presenter notes:</p>	

<div data-bbox="215 235 646 324" data-label="Section-Header"> <h3>Organisation should...</h3> </div> <ul data-bbox="239 347 678 481" style="list-style-type: none"> ▪ Allow you to find files easily ▪ Be meaningful to you and colleagues ▪ Be consistent 	<p data-bbox="829 235 917 264">Script:</p> <ul data-bbox="877 302 1388 873" style="list-style-type: none"> • In summary, no matter which type of data you are dealing with you need a system which will let you: • Find your files easily – researchers accumulate lots of material and your system should allow you to go straight to what you need. • Be meaningful – to both the original researcher and their colleagues. Mobile researchers often means a great deal of change. Good test of your system if your colleagues can use it. • Be consistent – choose a system and stick with it or it will just end up being more confusing.
<p data-bbox="204 638 411 667">Presenter notes:</p>	<p data-bbox="829 985 917 1014">Script:</p> <ul data-bbox="877 1052 1388 1310" style="list-style-type: none"> • Unless there is an existing local policy that you HAVE to use, your system can be whatever works for you. • Important thing is that the system allows you to find what you need when you need it and that other people can interpret it.
<div data-bbox="215 985 646 1075" data-label="Section-Header"> <h3>Organisation could...</h3> </div> <ul data-bbox="239 1097 718 1220" style="list-style-type: none"> ▪ Look however you want it to... ▪ ... as long as you (and others) can make sense of it <p data-bbox="204 1388 411 1417">Presenter notes:</p>	


<p>File naming conventions</p> <div data-bbox="236 338 772 443">  <p>My talk Meeting notes Results 20200602 Smith article</p> </div> <p>Would you know what any of these files were in three years?</p>	<p>Script:</p> <ul style="list-style-type: none"> • One of the easiest things to do is to use a file naming convention. • May seem like a hassle to set up initially but can save time and effort in the long run – would you know what any of these files were in three years time?
<p>Presenter notes:</p>	
<p>File naming conventions</p> <div data-bbox="245 1128 759 1301">  </div>	<p>Script:</p> <ul style="list-style-type: none"> • Many examples of systems out there – check if there is a local one. • One example from Queensland University of Technology uses a series of components to construct a file name: • Prefix – type of file. • Name – sensible name of the document. • Version number – which version/draft. • Date – of last update. • Person – who last updated it? • Files named using this system will display all types together, in order of version and when they were last updated = easy to find the most recent version of a file.
<p>Presenter notes:</p> <p>Animated slide – click to highlight each element of the convention.</p>	


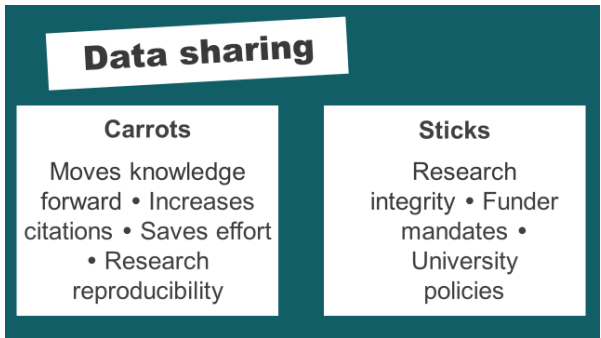
<div data-bbox="213 230 647 327" data-label="Section-Header"> <h2>File naming conventions</h2> </div> <div data-bbox="236 353 416 533" data-label="List-Group"> <ul style="list-style-type: none"> ▪ Objective ▪ Meaningful ▪ Concise ▪ Standardised </div> <div data-bbox="611 535 788 548" data-label="Text"> <p><small>Jisc Advice Guides: What Makes a Good File Name?</small></p> </div>	<div data-bbox="826 230 914 264" data-label="Section-Header"> <h3>Script:</h3> </div> <div data-bbox="874 304 1393 846" data-label="List-Group"> <ul style="list-style-type: none"> • Whichever system you use it needs to have the following characteristics: • Objective – a clear name that makes sense to everyone. • Meaningful – avoid abbreviations and personal shorthand. • Concise – long file names mean that the convenience of using a system is often lost and they can be harder to remember. • Standardised – system as a whole and individual components e.g. date format, should be standardised. • Guidance available from Jisc about file naming conventions. </div>
<div data-bbox="201 638 411 672" data-label="Section-Header"> <h3>Presenter notes:</h3> </div>	<div data-bbox="204 981 805 1317" data-label="Image"> </div> <div data-bbox="826 985 914 1019" data-label="Section-Header"> <h3>Script:</h3> </div> <div data-bbox="874 1059 1361 1339" data-label="List-Group"> <ul style="list-style-type: none"> • Once you have data organised you need to think about the next puzzle piece – how are you going to store the information you’ve used or created? • Most people just say that they will use a backup drive but there is still more to think about. </div>
<div data-bbox="201 1391 411 1424" data-label="Section-Header"> <h3>Presenter notes:</h3> </div>	



<div data-bbox="204 226 807 566"> <h3>Data disasters</h3> <div> <p>"King's College University fire: students evacuated after blaze in the Strand"</p> <p>- London Standard (2017)</p> </div> <div> <p>"Data-storage service's crash sets back researchers"</p> <p>- Chronicle of Higher Education (2014)</p> </div> <div> <p>"Nottingham university fire destroys new multimillion pound chemistry building"</p> <p>- Daily Telegraph (2014)</p> </div> <div> <p>"Manchester cancer hospital fire may have destroyed vital research"</p> <p>- The Guardian (2017)</p> </div> </div>	<p>Script:</p> <ul style="list-style-type: none"> • Backing up data is something we all know we SHOULD do but also something many of us leave until later when we 'have time'. • This can have devastating consequences as the stories on this slide show = all real headlines. • Extreme examples but think about smaller examples – misplaced memory sticks, computer crashes – and think about the amount of data which could be lost in each situation.
<p>Presenter notes:</p>	
<div data-bbox="212 994 759 1308"> <h3>Data backup</h3> <p>Method, Frequency, Location, Amount</p> <p><small>Jogjawa MD Badsha Meah via IconsHub CCO</small></p> </div>	<p>Script:</p> <ul style="list-style-type: none"> • Back to the jigsaw metaphor – putting together the individual pieces to create a bespoke backup strategy. • Different projects will need different plans. • Method – how will it be backed up? Is there a local, automatic system or do you have to set something up? External drives have a cost implication. Choose two methods that will store the amount needed and keep separately. • Frequency – how often will it be backed up? The more important data are and the more they change the more frequent backups will need to be. • Location – where will data be stored? Are there any restrictions on this e.g. keeping it in a certain geographic location? • Amount – how much will be backed up each time? Does everything need to be backed up or only parts? Think about how much you could stand to lose.
<p>Presenter notes:</p>	



<div data-bbox="204 226 807 562"> <h2>Key steps</h2> <ul style="list-style-type: none"> ▪ Always back up data ▪ In at least two locations ▪ Keep these apart from each other  </div>	<p>Script:</p> <ul style="list-style-type: none"> • Main takeaway message = always backup your information, in at least two locations and keep these apart from each other. • Not much use to have a strategy and then keep everything in one backpack which gets stolen!
<p>Presenter notes:</p>	
<div data-bbox="215 999 780 1301"> <h2>Cloud services</h2> <ul style="list-style-type: none"> ▪ Always read the small print ▪ Investigate University solutions ▪ Be careful which data you store  <p><small>Cloud: Yarrick Lung via IconsHub CC0</small></p> </div>	<p>Script:</p> <ul style="list-style-type: none"> • Many researchers use cloud services but it is important to use caution. • Read terms and conditions before signing up – you need to know what you are agreeing to and that it meets funder requirements e.g. storing data in the EU. • University has a range of institutional accounts = greater level of protection and more storage. CAUTION: remember that you will lose access to these when you leave. • Be careful with storing sensitive information with these services – do they meet the requirements?
<p>Presenter notes:</p>	



	<p>Script:</p> <ul style="list-style-type: none"> • Next step = archiving data. • Important step to both preserve research and prepare it for sharing so that others can come along and build on work which has also been done.
<p>Presenter notes:</p>	
	<p>Script:</p> <ul style="list-style-type: none"> • General Data Protection Regulation came into force in 2018. • Changed the rules governing personal how data can be collected and stored. • Expanded the definition of personal data and increased penalty for data breaches. • Consent forms need to inform participants exactly how their information will be used.
<p>Presenter notes:</p>	

 <p>GDPR</p> <ul style="list-style-type: none"> ▪ Data breach notification ▪ Right to be forgotten ▪ Increased responsibility for the researcher and the organisation ▪ Financial penalties 	<p>Script:</p> <ul style="list-style-type: none"> • University must be notified of any data breaches within 72 hours. • Right to be forgotten means people can have their details removed from research. • Organisation and individual researcher now have increased liability for any breaches = important to protect personal data. • Minor issues = maximum E 10 million fine, major issues = maximum E 20 million fine. • Information Compliance website of the University offers more information.
<p>Presenter notes:</p>	
<p>Rules of personal data</p> <ul style="list-style-type: none"> ▪ Don't collect it unless you need it! ▪ Informed, open, written consent ▪ Anonymise ▪ Use managed repositories 	<p>Script:</p> <ul style="list-style-type: none"> • Some basic advice to follow to ensure you protect personal/sensitive data. • Only collect information you actually need – personal information covers many things which can be used to identify someone. Think about if you need to collect each individual piece of information. • Obtain informed, open, written consent – participants need to sign a document saying they understand the research and that they are happy for results to be shared in print/online e.g. via a repository. • Anonymise all data – including splitting datasets if needed. • Use managed repositories to share data – complies with the sharing rules but also manages to keep data protected.
<p>Presenter notes:</p>	

	<p>Script:</p> <ul style="list-style-type: none"> • Sharing data = part of wider move towards open research. • Ensures that outcomes of research are shared with those who are interested. • Allows knowledge to move forward.
<p>Presenter notes:</p>	
	<p>Script:</p> <ul style="list-style-type: none"> • Different motivations for sharing data = carrots and sticks. • Carrots: • Allowing others to use data helps them in their own work and helps move global knowledge forward. • Increases citations for data creators = another output to cite. • Stops people repeating the same experiments to create the same results – especially if something doesn't work. • Reproducibility – need to be able to check and verify results = need to see data. More on this next. • Sticks: • Helps preserve integrity of research and protects research from speculation. • Funders and universities often now mandate that data underpinning publications is shared.
<p>Presenter notes:</p>	

	<p>Script:</p> <ul style="list-style-type: none"> • Reproducibility crisis = one of the most important reasons to share data. • Many scientific studies are impossible to reproduce – 70% can't reproduce the work of others and 505 can't reproduce their own. • Need to be able to verify results of research as part of academic process. • Several recent cases of fraud. • Making data available allows others to do this and test your conclusions = improved science. • Can also protect from accusations of misconduct.
<p>Presenter notes:</p>	
 <ul style="list-style-type: none"> ▪ At time of publication (or earlier) ▪ Select a suitable repository ▪ Describe and licence your data ▪ Investigate potential costs 	<p>Script:</p> <ul style="list-style-type: none"> • Covered the <i>why</i> of sharing but what about the <i>how</i>? • Data should be shared at the time of publication or even before publication and linked to the published version. Can actually protect from data theft as this date stamps the data. • Choose a suitable repository – some funders specify a repository but if not you will have to select one. • Describe/licence data – think about how you would like others to use your data. Add metadata/ReadMe file and a suitable licence. • Investigate costs – there are cost implications to sharing data, especially when there are large amounts. This will need to be built into budgets.
<p>Presenter notes:</p>	

	<p>Script:</p> <ul style="list-style-type: none"> • Repositories preserve and share outputs including data. • Range of repositories available – general, format/subject specific, institutional. • Need to think about what grant specifies, what your peers use and what will be the best venue for your output. • Try to use just one/two for different formats = easier to keep track of statistics. • Re3data = showcases a range of repositories and offers guidance on how to select the best one for your needs.
<p>Presenter notes:</p>	
	<p>Script:</p> <ul style="list-style-type: none"> • Cambridge IR = Apollo. Well indexed by Google and largely free to use. • Takes all types of output including data. • Can link publication and datasets together in Apollo.
<p>Presenter notes:</p>	

<div data-bbox="220 241 497 286" data-label="Section-Header"> <h2>Licensing data</h2> </div> <div data-bbox="220 313 466 524" data-label="List-Group"> <ul style="list-style-type: none"> ▪ How can your data be used? ▪ Can people publish their own work based on your data? ▪ Can others make money from your data? </div> <div data-bbox="507 226 804 562" data-label="Figure">  <p>As open as possible and as closed as necessary</p> </div>	<div data-bbox="826 230 912 264" data-label="Section-Header"> <h3>Script:</h3> </div> <div data-bbox="874 302 1385 1059" data-label="List-Group"> <ul style="list-style-type: none"> • Need to consider how you want others to use your data when sharing it. • You are the creator = you have the control. • Find a licence which outlines how you want your data to be used including whether you will allow them to use it in their own work and/or make money from what they create with it. • Funders may specify a certain licence or level of access e.g. CC, Apache, General Public Licence. • Many people are understandably worried about this but important to remember that in many circumstances you control the licence and you make the choice. • Advice is to be as open as possible and as closed as necessary = happy medium. </div>
<div data-bbox="202 638 410 672" data-label="Section-Header"> <h3>Presenter notes:</h3> </div>	<div data-bbox="204 1128 804 1464" data-label="Image">  </div> <div data-bbox="826 1097 912 1131" data-label="Section-Header"> <h3>Script:</h3> </div> <div data-bbox="874 1169 1353 1276" data-label="List-Group"> <ul style="list-style-type: none"> • All of the elements discussed come together in one document = data management plan. </div> <div data-bbox="202 1541 410 1574" data-label="Section-Header"> <h3>Presenter notes:</h3> </div>

<div data-bbox="215 235 683 331" data-label="Section-Header"> <h2>Data management plans</h2> </div> <ul style="list-style-type: none"> ▪ Outline of how data will be managed ▪ Mandated by funders ▪ Expected by the University ▪ Living document 	<p>Script:</p> <ul style="list-style-type: none"> • DMPs = outline of how data will be managed and shared. • Increasingly mandated by funders and institutions as a way to show that recipients will be responsible with the money they get. • Cambridge is increasingly mandating that they be drawn up prior to the start of a project = part of good academic practice. • Remember – DMPs are living documents which should be updated throughout the length of a project.
<p>Presenter notes:</p>	<p>Script:</p> <ul style="list-style-type: none"> • Different funders have different mandates about what needs to be in a plan but some common areas are highlighted on the screen. • Sources of the data – whether you have collected it yourself or are using existing data (and any permissions) • Type of data – what format(s) will the data be in? • Backup – how will data be stored and protected during and after the project? • Sharing – how will data be shared in compliance with both GDPR and funder requirements? • Ethics – which ethical standards will be adhered to and how will you ensure these are met? • Responsibility – who will have named responsibility for the documentation and data? Is there a succession plan?
<div data-bbox="202 978 804 1314" data-label="Diagram"> <p>Data management plans</p> <ul style="list-style-type: none"> Sources Did you collect new data? • Are you using existing data? • How? • Why? Types What form are the data? • Which file formats? • Which metadata scheme? Backup How will data be kept safe? • Where will it be stored? • How much will it cost? Sharing Will data be shared? • Where will this happen? • If not, why not? • Does this meet funder and local requirements? Ethics How will personal data be protected? • Is there other sensitive information? • How does this impact sharing? Responsibility Who has overall control of the data? • What happens if they move on? </div> <p>Presenter notes:</p> <p>Interaction – distribute handouts.</p> <p>Interaction – remind attendees about DMP course/resources.</p>	



Script:

- Any questions?

Presenter notes:

Action – send slides/online handout to audience.

Action – hand out feedback forms.