

Analysis of UK repository platforms: who is using what and why?

George Bray (Robert Gordon University, ORCID 0000-0002-7925-3347)

Presented at the M25 Consortium's cpd25 event, "Repositories: where are we now, where do we need to be?" on the 21st January 2025.

Abstract

Primarily focusing on the UK Higher Education sector, I aim to investigate the variety of repository platforms currently in use based on available data from UKCORR. Through an analysis of the data, I will attempt to determine whether there are any particular trends - geographic, financial or otherwise - evident from the current spread of platforms across the sector. Additionally, I will draw on informal and anecdotal evidence from repository staff to begin exploring institutional decision-making behind switching repository platforms.

Presentation Notes

[INTRO - STARTS AT 0 MINS]

1. [SLIDE 1] Personal introduction (1 min / 1 MIN)
 - Hello and thank you for the invitation to speak at this event.
 - My name is George Bray and I'm the Repository and Metadata Assistant Librarian at Robert Gordon University (RGU), a relatively small institution in Aberdeen.
 - I've worked professionally as a librarian since 2015, and I've had experience of working with two different repository systems: DSpace and Worktribe. It might also be worth noting that Worktribe is a research information system as well as a repository, and I currently chair one of the special interest groups for the Worktribe user community.
 - I'll be talking today about the kinds of repository software in use across the UK, and exploring potential patterns or causes for usage and system selection.

[WHAT COVERED - STARTS AT 1 MIN]

2. [SLIDE 2] What will be covered (1 min / 2 MINS)

- In particular, I'm going to cover the following:
 - Firstly, I want to explain my data sources and provide a few disclaimers.
 - Secondly, I'll discuss some relatively broad findings about repository usage in general, including commercial vs open source repositories, and uptake of data repositories.
 - Thirdly, I'll present some more focused findings about repository usage by "universities" (i.e. institutions that have been granted "university" title, or those that operate like universities e.g. Glasgow School of Art and Scotland's Rural College). This will include comparisons based on various aspects.
 - Finally, I'll outline some observations on the kinds of factors that institutions seem to prioritise when deciding whether to switch repository platforms or whether to stay with their current repository solution.

[DATA SOURCES - STARTS AT 2 MINS]

3. [SLIDE 3] Data sources (*combined total 3 mins / 5 MINS*)

- Repository information (1 min / 3 MINS)
 - So, in terms of data sources, my main source of information for repository usage has been a spreadsheet that is publicly available from UKCORR. Specifically, I've been working from a copy that I saved on the 21st November 2024.
 - I chose UKCORR's list for a few reasons:
 - One very minor reason was that it was already available in spreadsheet format and, lacking in time as I am, I thought this would make it easier for me to work with.
 - Secondly, given the nature of UKCORR, this information is entirely UK-focused and it therefore seemed highly relevant to the audience for this talk.

- Finally, the UKCORR spreadsheet is community-maintained and is also meant to be a very low-effort way for institutions to document information about their repository systems. I therefore thought that it might be more comprehensive and up-to-date than some other sources, which require more effort from people.
- However, I did end up finding a few issues with the UKCORR data - there are several institutions that I noticed missing, and many entries that haven't been updated for upwards of 5 years. Nevertheless, I felt it might produce some interesting insights, so I continued using it and just made a few corrections when I spotted out-of-date information.
- A couple of other potential sources that might be used in future include Jisc's OpenDOAR database or the euroCRIS Directory of Research Information Systems, though the latter isn't repository-focused.
- I also used developer websites and Wikipedia to identify which systems were commercial or open source.

[INFO ABOUT INSTITUTIONS - STARTS AT 3 MINS]

- **[SLIDE 4]** Information about institutions (1 min / 4 MINS)
 - Information about institutions came from a variety of sources:
 - "Size" of institution is based on the number of full-time equivalent students recorded in the Times Higher Education World University Rankings. The number of research staff or number of repository records might be more directly relevant to repositories, but I felt that student numbers were a reasonable proxy for institutional size and therefore probably also finances.
 - "Research intensity" is taken directly from the Complete University Guide University League Tables. It claims to be a proportion of staff (between 0 and 1) who are engaged in high-quality research, but this probably needs to be checked in more detail before it could be considered particularly reliable.

- "Age" is based on the date when university title was granted, or if it hasn't been granted, the date when degree-awarding powers were conferred. This is mostly based on information from the Office for Students, but I've had to use Wikipedia or university websites for unlisted and non-English institutions. This is the most problematic metric I've used - for example, many long-established institutions have only recently become "universities". However, complex and varied institutional histories make it difficult to find a consistent alternative.
- "Institutional groups" is based on member lists for three specific groups: GuildHE, MillionPlus and the Russell Group. There may be other groups that are worth exploring in future.
- The last category, "Region", is based on the regional categorisation from UCAS' website.

[INFO ABOUT DECISIONS - STARTS AT 4 MINS]

- [SLIDE 5] Information about decision-making (1 min / 5 MINS)
 - Information about factors affecting institutional decisions around repository switching and retention came from a combination of browsing the archives of the UKCORR mailing list and from recalling conversations with colleagues at various institutions. I would have liked to do more rigorous data collection and analysis here, but I'm afraid I haven't had time to do so.
 - Finally, and on a related note, I'd just like to emphasise that the data analysis I've done has been produced inexpertly and with relatively little available time, and is based on data sources that are varyingly informal and problematic. Please therefore don't take my observations as "truth", but rather consider them a starting point for future thought and investigation.

[REPOSITORY OVERVIEW - STARTS AT 5 MINS]

4. [SLIDE 6] Repository systems in use (*combined total 8 mins / 13 MINS*)
 - So, with all that in mind, let's dive in. I'm going to start by sharing some general observations about the repository systems that are in use.

- [SLIDE 7] Non-data repositories (2.5 mins / 7.5 MINS):
 - Looking first of all at institutions' main repositories, i.e. not including any repositories that are exclusively for research data, information in the UKCORR spreadsheet covers 132 different institutions, of which 118 are "universities" and 14 are other organisations (mostly a mixture of research institutes, funding bodies, conservatoires and non-higher education GLAM organisations).
 - These 132 institutions have 134 non-data repositories, using 12 different systems.
 - Only two institutions are listed as having two different non-data repositories, and in both cases they have one DSpace repository and one Pure repository. Even though we know that the UKCORR spreadsheet is incomplete, this still suggests that it is a relatively uncommon practice for an institution to have two non-data repositories. It is also worth noting that both of the institutions involved are universities, which makes sense because my general understanding is that this scenario is most likely to occur when an institution wants a completely separate repository for its theses.
 - Returning to the bigger picture, EPrints is clearly the most commonly-used repository system across all institutions, with 59 total instances (of which 53 are universities and 6 are other organisations). Pure is the second most-common with 23 instances, but is used only by universities. Meanwhile, DSpace is third most-common with 20 instances (of which 17 are universities and 3 are other organisations).
 - Given that there are far more universities in the spreadsheet than there are other organisations, it is not surprising that there is a wider range of systems in use among universities. However, it might be worth noting that Hyku is used only by other organisations, not by universities. I don't have any information about why this might be, but one suggestion is that it might relate to the specific nature of the research conducted at the heritage organisations in question.

[COMMERCIAL VS OPEN - STARTS AT 7.5 MINS]

- [SLIDE 8] Commercial vs open source (1.5 mins / 9 MINS):
 - Looking further at the business models behind these non-data repositories, we can see that there is an equal split, with six commercial platforms (of which two are owned

by Elsevier!) and six open source platforms (of which two are developed by DuraSpace).

- However, when we translate this into the frequency with which systems are being used, there is a significant difference. Across all institutions, roughly a third of repositories use commercial systems versus two-thirds that are using open source systems.
- These proportions are fairly similar when looking only at universities, but there is a slightly more pronounced bias in favour of open source systems when looking only at other organisations.
- It could be argued that the slight bias among other organisations in favour of open source solutions might be related to finances. Repositories might play a smaller role in the overall activities of these organisations and therefore be less-well funded than repository services at universities, and open source solutions are typically cheaper to get up and running than buying a licence for a commercial product.
- I don't have authoritative evidence to explain the high usage of open source solutions in general, but some of the factors involved in institutions' decisions will be explored later. However, it is worth noting at this stage many institutions' commitment to open infrastructure, which makes sense given that repository services by their nature are generally driven by a desire to make things open.

[DATA REPOS - STARTS AT 9 MINS]

- **[SLIDE 9]** Data repositories (3.5 mins / 12.5 MINS):
 - So far I've just spoken about institutions' main repositories, but what about data repositories? In the UKCORR spreadsheet, most institutions (both universities and other organisations) do not have an entry for their data repository. In some cases this is because the data is incomplete, but it might also suggest that Open Data is a lower priority for many institutions in comparison to Open Access for other research outputs (e.g. articles and papers).
 - The chart at the top-left shows that just under half of all institutions have an entry for their data repository and only one of these is a non-university organisation. This is

perhaps because there is greater emphasis on data sharing for university researchers, rather than researchers elsewhere. The chart also shows that, where institutions **have** reported a data repository, there are three different scenarios:

- Firstly, where the institution uses the same repository instance for both data and non-data;
 - Secondly, where the institution uses separate repositories for data and non-data, using different software;
 - And thirdly, where the institution uses separate repositories for data and non-data, but using the same software for both.
- The chart at the top-right takes these three scenarios and maps them to the systems used for the institution's non-data repository. There are a few observations here:
- Interestingly, relatively few EPrints institutions use a single repository for data and non-data, whereas Pure and Figshare are much more prominent in this regard. Figshare was designed primarily as a data repository, so it makes sense that most of the institutions who are using it for their main repository are also using the same repository for their data. As for Pure, the rationale here is less clear, but perhaps it relates to the fact that Pure is both a repository and a research information system; however, almost as many Pure-using institutions are using a different system for their data repository, so this isn't a particularly strong argument.
 - Only two systems appear for the scenario where an institution has separate data and non-data repositories using the same software - DSpace and EPrints. EPrints is far more common in this regard, whereas institutions that use DSpace for their main repository are more equally split across different scenarios for their data repository. Again, there's no clear reasoning as to why it is so common for EPrints-using institutions to have two separate instances of it, one for data and one for non-data. One potential factor is that institutions may find it easier to customise EPrints to optimise the way in which their data is shared. However, institutions with EPrints as their main repository are

almost as likely to use different software for their data repository, so again this isn't a strong argument.

- Finally, the table at the bottom explores the scenario where an institution has separate data and non-data repositories using different software. Given its origins as a data-focused repository, it is perhaps unsurprising that Figshare is so common here, regardless of which non-data repository software the institution uses. However, it is by no means the only data repository solution and we can see a relatively wide variety of systems being used here, including two (Zenodo and Sufia) that aren't present in our list of systems for non-data repositories.

[UNI FOCUS - STARTS AT 12.5 MINS]

5. [SLIDE 10] Focus on "university" and non-data repositories (*combined total 14.5 mins / 27 MINS*)
 - There's scope for a lot more analysis and theorising around data repositories, especially given the different needs that institutions might have for their data in comparison to other research outputs. However, I'm going to leave that for future investigators to explore and will now move on to focusing on non-data repositories again, this time looking exclusively at those used by universities, so that I can more easily compare them in various ways.

[BY SIZE - STARTS AT 13 MINS]

- [SLIDE 11] Repositories by size of institution (1.5 mins / 14.5 MINS):
 - I'm firstly going to compare university repositories based on size of the institution (i.e. number of students).
 - EPrints is the most common system across all size categories, and with roughly equal numbers of instances in each category. However, it is especially prominent among the "very small", "small" and "very large" categories, where it is used by roughly 73%, 50% and 50% of institutions in each, respectively. In comparison, it is used for only around 25-30% of repositories in the middle categories. These middle categories also feature the largest number of institutions though, so perhaps it makes sense for a greater variety of repositories to be found here and therefore for EPrints to have less of a market-share.

- For the other two main systems, we can see that Pure is also equally common across all size categories, with particularly high representation in the "small", "high medium" and "very large" categories where it represents about 25% of repositories in each. Meanwhile, DSpace is used for 53% of repositories in the "low medium" category, but is less common elsewhere.
- Worktribe is the only other system that is present across more than two different size categories. Now, there isn't enough evidence to make deductions about the offerings of other systems, but the available information does suggest that at least EPrints, Pure, DSpace and Worktribe are capable of meeting the needs of many UK institutions regardless of their size.

[BY RESEARCH INTENSITY - STARTS AT 14.5 MINS]

- [SLIDE 12] Repositories by "research intensity" (*2 mins / 16.5 MINS*):
 - Looking now at "research intensity", I've grouped institutions into three categories: "high", "middle" and "low". (Data wasn't available for seven institutions, so I've excluded them here.)
 - Interestingly, EPrints is less common than either DSpace or Pure among the most research-intensive institutions. However, it vastly outnumbers other systems in the "middle" and "low" categories. I'm not sure why; regarding Pure, it might relate to institutions in the "high" category having more research income and therefore being better positioned to purchase commercial solutions, but DSpace is just as popular in this category as Pure and it is open source, so that doesn't really hold up. It might be that it costs more to host a DSpace repository than an EPrints one, but that theory would require further investigation.
 - We can also see that both Pure and DSpace are significantly less common in the lowest category than they are in the two higher categories. This might be related to costs, as I just mentioned, or perhaps it might be that these systems provide more functionality than is required by less research-intensive institutions. However, all we can really say is that there doesn't seem to be a strong correlation between research intensity and institutional preferences for commercial vs open source platforms.

- Figshare and Worktribe are the only other systems that are represented in all three categories, with fairly equal distribution across each. Finally, it is worth noting that the "middle" category features the greatest variety of systems, which is reasonable considering that it is also the category with the greatest number of repositories.
- Again, while the available information does support the fact that EPrints, Pure, DSpace, Worktribe and Figshare are able to meet the needs of many institutions regardless of research intensity, there isn't enough evidence to suggest the opposite for other systems.

[BY AGE - STARTS AT 16.5 MINS]

- [SLIDE 13] Repositories by "age" of institution (2 mins / 18.5 MINS):
 - For age, I've grouped institutions into the following categories: "Ancient", "18th/19th century", "1900-1959", "1960-1991", "1990s" (i.e. '92-'99), and "2000+".
 - One observation here is that there are a relatively small number of systems in use among the older categories, but these categories are actually still quite varied because this small number of systems covers a very small number of repositories.
 - It's also interesting to note that DSpace is more popular than EPrints among "Ancient" universities, but isn't used at all in the "18th/19th Century" category, where EPrints again returns to being the most common platform. Institutions from the first half of the 20th century show a clear preference for EPrints, which is a trend that continues to increase with each more recent age category, strongly implying that the newer the institution, the more likely it is to use EPrints.
 - Only EPrints is present across all age categories, but Pure, DSpace, Worktribe and Figshare are found across most categories, so the data suggests that all five systems are suitable for institutions regardless of age.
 - A final couple of observations here:
 - Firstly, the only two instances of an institution having multiple non-data repositories are both in the "Ancient" category, which suggests that older

institutions might have different attitudes towards how theses are dealt with on their repositories.

- Secondly, Haplo (now Cayuse) is used only by 1990s universities - in fact, it is the third most common platform for institutions in that category, rather than DSpace. It's unclear why this might be, but it is interesting to see, nevertheless.

[BY INSTITUTIONAL GROUP - STARTS AT 18.5 MINS]

- [SLIDE 14] Repositories by institutional groups (2.5 mins / 21 MINS):
 - I thought it might be worth comparing institutions based on membership of certain groups, specifically GuildHE, MillionPlus and the Russell Group, due to the similarities that might exist between members of each group.
 - GuildHE says that its members are principally focused on vocational and technical higher education.
 - MillionPlus says that its members are "modern" universities, though it is unclear from its website as to exactly what that means.
 - Meanwhile, the Russell Group are supposedly some of the most research-active universities in the UK, and - according to Wikipedia at least - the combined membership of this group receive over three-quarters of all university research grant and contract income in the UK.
 - I should also note that many institutions aren't in any of these groups, so this comparison is just based on repositories for 46 institutions.
 - So one immediate observation is that there is far more system variety for MillionPlus and Russell Group institutions, since almost all GuildHE institutions use EPrints. This is probably due to GuildHE offering a shared repository service - in case you're not sure what that means, it is basically a central organisation hosting different repository instances for multiple institutions, all using the same software.
 - EPrints, Pure and DSpace are the most common repositories across MillionPlus and Russell Group institutions. However, while proportions of Pure and DSpace users are

roughly the same across both groups, EPrints is particularly prominent among Russell Group universities, being used in 46% of repositories there.

- Finally, the numbers are very low, but it may also be worth noting that - of the 46 institutions represented here - the only instances of Haplo and Esploro are found for MillionPlus institutions, and the only instances of Worktribe and Fedora are found for Russell Group institutions. However, I don't think there's sufficient information here to suggest that any of those repositories is better suited for "modern" or highly research-funded institutions. On the other hand, there is a slight suggestion that MillionPlus members might be more likely to use commercial platforms, as these account for 40% of repositories in that group, versus only 33% in the Russell Group. This contradicts some of my earlier thoughts about how higher research income might correlate with a greater number of commercial repositories.

[BY REGION - STARTS AT 21 MINS]

- [SLIDE 15] Repositories by region of the UK (3 mins / 24 MINS):
 - The final aspect I want to compare is geographical region and in fact there are some interesting observations here:
 - Firstly, while EPrints remains the most common repository system in several regions, this is most strongly the case in London, "Yorkshire and the Humber", and several western regions of England. In most other regions (including the three spanning the majority of eastern England), it is either as popular as - or even less popular than - one or more other repository systems. I also find it particularly interesting that EPrints usage is mostly an English phenomenon, being significantly less popular in Northern Ireland, Scotland and Wales.
 - Secondly, while Pure is generally represented across the majority of regions, it is particularly common in Scotland and South-East England.
 - Thirdly, DSpace is relatively rare in the southern parts of England, and it is further interesting to note that three of the four instances there are in London. However, it is much more prominent in Scotland and the East of England region.

- My last main observation and something I find particularly interesting is that, although London has the most repository instances out of all the regions (22 repositories), it also has only four different systems in play, giving it the lowest figure in terms of variety of systems when considered against the number of repositories.
- Beyond these more substantial patterns, there are a few other things that might be worth noting as well:
 - I previously noted that there were only two instances where an institution had multiple non-data repositories, and that in both cases the institutions were using DSpace for one repository and Pure for the other. We can see here that both instances are actually in Scotland and that this coincides with Scotland having a higher number of DSpace and Pure repositories in general, when compared to other regions.
 - A second minor observation is that all five instances of Haplo are in London or South-East England, and I find it really interesting that a system with so many different institutions using it is so limited in geographical spread. This is in contrast to Figshare and Worktribe, which are both in use at a similar number of institutions as Haplo, but for which the institutions involved are dotted around the country.
 - Finally, it is interesting that several of the least-commonly used systems are all found in South-East England, which might suggest that institutions in that region are perhaps more open to using systems that are less mainstream.
- The general impression from these observations is that there might be some correlation between region and repository system. The most obvious explanation for this is that institutions are likely to look at their nearest neighbours when considering which repository system to use themselves.

[VARIETY SCORE - STARTS AT 24 MINS]

- [SLIDE 16] "Variety scores" (*3 mins / 27 MINS*):
 - After having made these comparisons based on different institutional aspects, my observation about the relative lack of variety in systems in London got me thinking about some way of comparing all these different aspects on the basis of system variety.
 - Essentially, what I've done is divided the number of different systems by the total number of repositories for each category, to produce what I've decided to refer to as a "variety score" - a number between 0 and 1. The closer a score is to 1, the greater the variety of repository systems seen for that category, and the closer a score is to 0, the less variety there is in that category.
 - It doesn't really work as a metric for small categories. For example, we know that there's only one system used in Northern Ireland, but because there are only two repositories there, the "variety score" is 0.5, which is actually rather high when compared to other categories.
 - Small categories aside, the majority of categories have a "variety score" lower than 0.5, implying that there is a relatively low amount of variety in repository systems being used in the UK in general. However, there are still a few potential trends that can be observed:
 - Firstly, younger universities are more likely to have the same repository software as each other, whereas there is more likely to be variety among older universities. However, this is potentially contradicted by the fact that membership of the "modern" MillionPlus group appears to coincide with a higher variety score than membership of the Russell Group, which might be assumed to feature older universities. My suspicion is that this just reflects the problematic nature of "age" as a category, as I've mentioned previously.
 - Secondly, research intensity doesn't seem to have very much impact on system variety, as the scores for all three categories in this area are quite close together. However, size of institution does seem to have more of an impact, with there being less variety among "very small" and "very large" institutions,

in comparison to much greater variety among institutions in the middle size categories.

- The third trend here is that geographical region might be argued to have relevance to system variety, because some regions are much more varied than others. For example, I've already mentioned that London has the highest number of repositories but the lowest variety; however, the regions that are second and third in terms of numbers of repositories - i.e. South-East England and Scotland - actually have quite high variety scores in comparison. Then there are also some regions that dominate the upper end of this chart with significantly high variety scores, especially the East Midlands, North-East England and Wales - and it's worth noting that each of the five repositories in Wales uses an entirely different system. However, on the whole, most region categories have variety scores that are lower than 0.5, which again supports the suggestion I made previously that institutions are relatively likely to be influenced by the repository choice of their neighbours.

[DECISION-MAKING - STARTS AT 27 MINS]

6. [SLIDE 17] Why switch? Why stay the same? (*combined total 3 mins / 30 MINS*)
 - So, on the whole there is relatively little to suggest clear patterns or trends in software usage based on the quantitative data. I want to conclude this presentation by highlighting a few observations from qualitative data, regarding how institutions have talked about deciding whether to switch repository systems or to keep their existing ones.
 - [SLIDE 18] Example of decision-making criteria (1.5 mins / 28.5 MINS):
 - One example is from White Rose Libraries, who made a recent announcement about their decision to retain EPrints. (I've included a screenshot of it, but this is just for illustration.) I think this is a good example because it echoes many of the factors that I've seen brought up elsewhere.
 - The first thing they say when they start talking about decision-making criteria is that they wanted to evaluate "[...] the costs and value [...]" of the options available.

- I think it is noteworthy that costs are mentioned first, though it isn't until later in the announcement that further detail is provided, where they mention high costs of commercial products and the costs of accessing technical expertise for open source solutions. They also mention being limited by difficult financial circumstances.
- The "value" part is less clearly defined, but there are various statements that are relevant, for example about the importance of being able to customise the repository, ensure that it meets UK-specific needs, and can support White Rose Libraries' shared service set-up. They also mention functionality and the importance of a thriving user community.
- In addition to costs and value, technical or logistical considerations are mentioned, such as how easy it is to migrate to a new platform, reliability of service levels and long-term sustainability.
- Finally, White Rose Libraries say that they have a "long-standing commitment to open access and to open source infrastructure", which played a key role in their preference for open source solutions.

[MAIN FACTORS - STARTS AT 28.5 MINS]

- **[SLIDE 19]** Main factors (1.5 mins / 30 MINS):
 - Based on what I've seen on mailing lists and what I've heard during conversations with colleagues at various institutions, the factors affecting repository decision-making seem ultimately to boil down to four main areas:
 - Cost is always a factor, especially given the financial issues currently affecting many institutions. There is the cost of the infrastructure - licences for commercial solutions, or server costs for open source solutions. Then there are also costs associated with human resources - for example, needing to ensure that you have access to requisite technical expertise if you are hosting an open source repository. For institutions looking to cut costs, a new repository might also appeal if it can offer more streamlined workflows that enable an institution to reduce the size of its repository team. Cost is perhaps a more

decisive factor for smaller institutions than larger ones, but no one can ignore it.

- While not as consistently important as cost, timing can also be very important, as a repository migration needs to be considered alongside other events. These might be at sector level - for example an institution is unlikely to change repository just before their REF submission, or to invest in a big software change during times of high economic uncertainty. There may also be local events to consider, such as other IT projects, or fluctuations in staffing and available skill sets - for example, if your local DSpace expert retires.
- Functionality is also really important. Most repositories will do the "basics", but institutions are likely to evaluate options for more advanced functionality, such as support for research data or the ability to integrate with other systems. Institutions also value how well the system meets local needs, for example being able to respond to UKRI's technical requirements or the new REF Open Access policy, or systems that enable an institution to use its own branding or adapt the user interface. A particular example of the latter are arts-focused universities, who want to make their repository content as visually-engaging as possible.
- Finally, there is the importance of community. Institutions want to see that their system is being actively developed, which requires an active user base. Institutions also value access to peer support networks, especially for discussing UK-specific needs or issues.

7. [SLIDE 20] Thank you for listening and I hope you've found my talk interesting. I plan to make my slides, notes and the spreadsheet of quantitative data available on my institution's repository and will share a link once I've done so.

[ENDS AT 30 MINS]

Speaker Bio

I completed my bachelor's degree at the University of Oxford, then some years later worked as Graduate Trainee Library Assistant at the Institute of Advanced Legal Studies. While studying for my master's in library

and information studies at University College London, I volunteered at the libraries of the Worshipful Company of Brewers and the British Museum (Greece and Rome). My first professional role was maternity cover at Cranfield University library, where I was involved in their services for interlibrary loans and Open Access.

I became Repository and Metadata Assistant Librarian at Robert Gordon University (Aberdeen) in 2016. Although I still have the same job title, my responsibilities have changed quite a lot over the years, covering things such as cataloguing (now mostly done elsewhere in the team), running our open journal platform, and supporting researchers with things such as Open Access, bibliometrics and outputs-based reporting. I have chaired the special interest group for the Outputs module of the Worktribe CRIS since 2023.

I am particularly interested in metadata and Open Access. I am also a keen advocate of games-based learning, having co-organised several "Research Support Games Days" alongside Valerie McCutcheon from the University of Glasgow.