Open to Open Access 'O2OA'

Needs Assessment summary (final)

REVISED January 2015

JISC Pathfinder Project team:

Coventry University:	Dr Lorna Everall (PI), Julie Bayley, Joanne Marsh.
University of Northampton**:**	Miggie Pickton, Nick Dimmock, Katie Jones.
De Montfort University:	Alan Cope, Alan Brine, Katie Fraser.

Contact:	Julie Bayley (Project manager)
Email:	j.bayley@coventry.ac.uk
Project blog:	http://blogs.coventry.ac.uk/researchblog/category/oa/
Document version:	Version 3 (*Final plus revised number of UCL Pre payment plans, pg 11*)

1) Report summary

This report summarises the findings of the O2OA Needs Assessment (Work Package 1). The aim of WP1 was to understand the OA needs of the academic and research support communities, both for publications and data, to form the basis of workflows and processes developed in WP2.

Acknowledgements

The team would like to thank all participants across the institutions, and those team members who undertook data collection. We would also like to thank partners from other Pathfinder projects who have shared learning from their activities.

Contents

2) Project and Team Summary

O2OA is a joint JISC funded project between Coventry University, the University of Northampton and De Montfort University. These three Midlands-based, Modern Universities are collaborating to identify issues and challenges – and ultimately recommendations – for the implementation of OA requirements in a modern university setting.

2.1 Aims, Objectives and Final Output(s) of the project

The overall aim is to establish shared institutional processes for the HE sector to support Open Access data management, OA publications and associated information management processes. More specifically the project aims to:

- Identify the **OA needs** of academics, information managers, research support staff, corporate leads and external funders
- Understand the **relationships** between OA publications, OA data and impact
- Develop **workflows** for OA processes
- Test and review **technical modifications** to existing institutional systems
- Embed a **pro-OA culture** using behaviour-change informed approaches to engage academics
- Align recommendations and workflows to funders' and institutional **mandates**
- **Disseminate** this learning across the HE sector

Throughout the project we will produce and disseminate a range of outputs to share the learning sector-wide such as:.

- Needs assessment report (October 2014)
- Agreed set of OA workflows and processes (October 2015)
- Posters, webinars, blogs, conference presentations and workshops (ongoing)
- Good practice handbook (May 2016)
- JISC project workshop (spring 2016)
- Ongoing social media (eg. Twitter, LinkedIn, Blogs)

2.2 Wider Benefits to Sector & Achievements for Host Institution

Benefits will include:

- **Increased awareness and knowledge** of OA and associated processes among stakeholders (researchers, research managers and administrators, senior management).
- **Increased compliance, adoption and engagement** of OA policies
- **Improved workflows** and processes for OA
- **Reduced burden** on HEIs to implement OA

2.3 Project Team

The collaboration draws on the expertise and experience of each partner to reflect the diversity of experiences within the HE sector:

Coventry University (CU, lead organisation):
Dr Lorna Everall (project director), Julie Bayley, Joanne Marsh.

Coventry leads on the overall project and on OA data management, drawing on their expertise in impact system development, embedding impact institutionally and effecting culture change within academia. Coventry are currently auditing current research data management practices and will develop associated policies and procedures

University of Northampton (UoN)
Miggie Pickton, Nick Dimmock, Katie Jones.

Northampton leads on OA Publications, having a particular interest in increasing its proportion of OA publications in line with the recent HEFCE policy for open access in post-2014 Research Excellence Frameworks.

De Montfort University (DMU)
Alan Cope, Alan Brine, Katie Fraser.

DMU leads on aligning OA with CRIS systems. DMU is in the process of implementing a new CRIS system and will develop associated workflows and process improvements. This will improve DMU's research support systems and institutional aim of increasing Open Access publications.

2.4 Timescale

The project runs from May 2014-May 2016. Work begins with a needs assessment (June-Oct 14), followed by process and workflow development (Oct 14-Sep 15), technical review and refinement (Apr 15 - Feb 16) and culminating in a range of tools and guidance in May 2016.

3) Needs Assessment Summary

The aim of WP1 was to understand the OA needs of the academic and research support communities, both for publications and data. Given the nascency of OA mandates, particularly for data, a qualitative approach was used to elicit beliefs, attitudes and understanding of stakeholders. This, rather than a quantitative approach, supports the development of user-led, ground up workflows and processes in WP2.

Specifically the assessment sought to explore:

For OA Publishing:

1) Drivers for the university (e.g. REF, funders)
2) Knowledge about publishing routes (Gold vs. Green, role of institutional repository)
3) Main reasons for publishing via an OA route
4) Reservations about publishing via an OA route
5) Confidence in having sufficient knowledge to choose the most appropriate route
6) Institutional support and facilitators

For OA Data:

1) Drivers for OA data (e.g. REF, funders)
2) Knowledge about OA data
3) Main reasons for making data open
4) Reservations about making data open
5) Confidence in having sufficient knowledge to comply with OA data guidance
6) Institutional support and facilitators

Originally intended for delivery at the end of September 2014, the process of data collection proved far more complex and time intensive than expected. Primary complications included:

- Difficulty recruiting over the summer period
- Staffing changes within institutions (both for gatekeepers and participants)
- Institutional strategy / changes
- Difficulties elevating the priority of OA to recruit participants

This report first summarises the context of OA support and compliance, including publisher resources, followed by a detailed analysis of qualitative data. Translations of findings into recommendations for workflows are also given.

4) OA context

4.1 OA background and rationale

The traditional model of publishing presents significant barriers for access within and beyond the academic community, and the speed at which information is disseminated. The timescale of post-review print publishing and more significantly the permissions associated with journal subscriptions has left much scientific output locked behind paywalls and inaccessible to the broader community. For many funders, academics and non-academics alike, these restrictions prohibited timely uptake of findings and in places prevented key audiences from access.

The Open Access agenda has grown to address these difficulties and increase availability of findings in the broadest terms; OA's key premise is free access *for all* to enhance uptake and use of findings. In parallel technological advancements in recent years have allowed the sector to shift away from printed (only) journals and capitalise on the viability of online access to e-articles. Thus both conceptual and technical developments have converged to offer a user-driven platform for data release.

The most fundamental change to facilitate OA is a change in the publication costing model. Traditionally authors published for free, with publication costs funded through reader subscriptions and one-off access charges. In a full turnaround, the OA model requires authors to pay to publish, allowing readers to have free access and removing subscription only access. Operationally this model allows both full 'Gold' Open Access (wherein authors pay a fee for their article to be released immediately by the journal) alongside a free 'Green' option wherein publishing is free but subject to embargo periods and deposit on institutional repositories. These two routes maintain free access to users but offer feasible options for authors who cannot afford a full Gold fee.

The upshot of these moves is a series of key benefits for authors, institutions and users alike. These include:

- Immediate and free access to articles for academics, non academics (inc. businesses, public, schools and colleges, practitioners)
- Increased visibility of academic findings, within and across disciplines
- Increased discoverability, crucial to academic reuse and broader impact
- Improved opportunity to find and use information for academic and non academic purposes
- Speeds up the process of scientific advancement by accelerating the release of information
- Maximises value from funded activities by amplifying the reach of results

Thus as the sector innovates its publishing model, OA offers an inclusive and viable means to disseminate academic advancements broadly and effectively.

The rest of this section headlines some of the key information and formal drivers for OA, along with some key information and links to external sources for further information.

4.2 HEFCE REF2020 policy

To be eligible for submission to the next REF (likely 2020), authors' final peer-reviewed manuscripts must have been deposited in an institutional or subject repository on acceptance for publication. Deposited material should be discoverable, and free to read and download, for anyone with an internet connection.

The requirement applies only to journal articles and conference proceedings with an International Standard Serial Number. It will not apply to monographs, book chapters, other long-form publications, working papers, creative or practice-based research outputs, or data. The policy applies to research outputs accepted for publication after 1 April 2016, but we would strongly urge institutions to implement it now.

4.3 Funder policies

SHERPA Juliet provides summaries of funding agencies' grant conditions on self-archiving of research publications and data. As a tool it offers considerable intelligence on funders' mandates and what archiving is allowed.

Research Councils UK (RCUK)
RCUK covers the following councils:

- Arts and Humanities Research Council (AHRC)
- Biotechnology and Biological Sciences Research Council (BBSRC)
- Engineering and Physical Sciences Research Council (EPSRC)
- Economic and Social Research Council (ESRC)
- Medical Research Council (MRC)
- Natural Environment Research Council (NERC)
- Science and Technology Facilities Council (STFC)

Researchers are expected to publish in journals that are compliant with the RCUK policy on Open Access (including CC-BY licence), and to acknowledge the receipt of Research Council funding. The policy applies to all peer□reviewed research articles which arise from Research Council funding, that are submitted for publication from 1 April 2013, and which are published in journals or conference proceedings. RCUK will allow some flexibility in implementation of its policy, including regarding the length of embargo periods, during the transition period (2013-18).

National Institute for Health Research (NIHR)
NIHR set out a clear commitment to OA and give a clear set of statements on expectations:

- The Department of Health and the NIHR require that, NIHR funded researchers seek to publish their research outputs (as outlined in the scope) in a peer-reviewed journal that is compliant with the policy on Open Access

- Where NIHR funds are used to pay an Article Processing Charge (APC) it is expected that the publication should be published in a journal that makes the output available using the Creative Commons Attribution (CC BY) licence, and allows immediate deposit of the final published version in other repositories without restriction on re-use. If this is not possible then the authors must reconsider the appropriateness of the original publication and how to comply with the point above.

  From the commencement of this policy (see above) the cost of APCs and other appropriate publication charges will be supported either through the original NIHR award, or where applicable an NIHR Open Access Fund. Researchers will need to contact the awarding NIHR Programme Coordinating Centre to review the correct approach. This does not apply to any recipient of NIHR funding prior to 1st April 2014, which remain under the requirements of their original contract.

- None of the above should prevent NIHR funded authors from also depositing a copy in their own institutional or another subject-based repository should they choose to do so or be required to do so by their employing institution subject to any restrictions from the publishers.

Charities
Most charities OA have similar features asking for funded research to be made available either through Gold OA or via Green OA. Some charities ask that a copy is deposited in PubMed Central (PMC) or Europe PubMed Central (Europe PMC). Most stipulate a maximum embargo period ranging from 6 to 12 months. Some examples of Charities OA policies:

- Arthritis UK
- Breast Cancer Campaign
- British Heart Foundation
- Cancer Research UK
- Wellcome Trust

Funder's data policies

RCUK, Cancer Research UK and Wellcome Trust outline their data deposit policies, with the following summary table (figure 1) providing an overview of the coverage of each funders' stipulations:

Figure 1: Summary table of funders' OA policy.

● Full Coverage ◐ Partial Coverage ○ No Coverage

	Policy Coverage		Policy Stipulations					Support Provided			
Research Funders	Published outputs	Data	Time limits	Data plan	Access/ sharing	Long-term curation	Monitoring	Guidance	Repository	Data centre	Costs
AHRC	●	●	●	●	●	◐	○	●	○	◐	◐
BBSRC	●	●	●	●	●	●	●	●	●	◐	●
CRUK	●	●	●	●	●	●	●	◐	●	○	○
EPSRC	●	●	●	◐	●	●	●	◐	○	○	●
ESRC	●	●	●	●	●	●	●	●	●	●	◐
MRC	●	●	●	●	●	●	○	◐	●	○	◐
NERC	●	●	●	●	●	●	●	●	●	●	◐
STFC	●	●	●	●	●	●	●	◐	●	◐	◐
Wellcome Trust	●	●	●	●	●	●	●	●	●	◐	●

Taken from http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies

4.4 Publishers and OA self archiving

SHERPA/RoMEO: is a searchable database of publisher policies on self-archiving of papers (i.e. *green open access*). It also contains links to information about publishers' compliance with funder policies (also covered by SHERPA/FACT), and their pages on copyright and paid for open access (OA) options.

Some RoMEO statistics (accessed 18/8/14):

- 575 publishers allow authors to deposit the publisher version or PDF of their article in an Institutional Repository, without fee or an embargo and a further 22 after publisher permission has been obtained (source).
- 67 publishers allow authors to deposit the publisher version or PDF of their article in an Institutional Repository after an embargo period of up to one year, a further 24 after a longer embargo period (source).
- Total: **74%** of publishers on the RoMEO database formally allow some form of self-archiving:

RoMEO colour	Archiving policy	Publishers	%
green	Can archive pre-print and post-print	556	34
blue	Can archive post-print (ie final draft post-refereeing)	550	33
yellow	Can archive pre-print (ie pre-refereeing)	112	7
white	Archiving not formally supported	437	26

(source)

RoMEO also has a list of publishers that have paid for open access options (source).

4.5 Article Processing Charges:

Commercial publishers have enthusiastically embraced *gold open access* since, by charging article processing charges (APCs), they gain another revenue stream. Many publishers have turned their journals into 'hybrid' journals comprising a mix of OA content (if the author pays an APC) and subscription content (for those who can't or won't pay the APC).

The size of APCs varies enormously between publishers and even between journals. Some publishers (e.g. PLoS) offer reduced or zero fees for those researchers who would otherwise be unable to publish and others (e.g. Taylor & Francis) offer discounted deals to institutions on payment of a fixed annual sum.Some universities have bought into these deals substantially, reasoning that their researchers are publishing sufficient papers with the publishers to justify the up front expense (UCL have 12 such pre-payment agreements) (see Jean Harris's presentation at Repository Fringe) and that the deals will save them time too.

SHERPA/RoMEO is a good starting point for establishing if a publisher offers an OA option. For each RoMEO journal record there is a link to the publisher's copyright policy and paid OA options.

There are other, mostly newer, publishers which have jumped on the bandwagon, creating in some cases dozens of new scholarly journals. Some of these publishers are quite aggressive in contacting academics to write for them. Some have dubious peer review processes and there are horror stories of respected researchers' names being used on fake Editorial Boards and the like. (See this blog post for more information). Even where the publisher is reputable there is some concern about whether publishers are actually providing the type of OA that has been paid for and whether this is the full OA (CC-BY) which some funders (e.g. RCUK) expect. The Jisc Monitor project is exploring how a JISC managed shared service might support institution in monitoring their APC spend, as well as what has been published within an institution and its compliance with funders policies.

4.6 Jisc: Open access and publishers

Jisc has produced a useful summary of the issues involved in gold open access publishing including issues such as the time and costs of handling APCs, the allocation of research funding to pay APCs; and the need for transparency (especially to counter claims of 'double-dipping'). As part of the JISC Total Cost of Ownership project, the Data Capture and Process report highlights the growth and management of APCs by HEIs. The data collected for this will underpin work to model APC offsetting schemes and support JISC Collections in negotiating with publishers. Jisc Collections has also, extended its remit from journal subscriptions only to now include APC payments; negotiating deals with publishers and investigating the efficiencies of JISC collections involvement in establishing processes, workflows and systems to subscriptions and payments. A particular concern at the moment is whether payment of APCs can be offset against subscription costs (to avoid the aforementioned double dipping). July 2014's JISC Watch from SCONUL (an informal newsletter) has an update on this.

4.7 Publishers Association:

The Publishers Association is the leading trade organisation serving book, journal, audio and electronic publishers in the UK. Membership comprises 117 companies from across the trade, academic and education sectors. Its core service is representation and lobbying, around copyright, rights and other matters relevant to members.

The Publishers Association welcomed the recommendations of the Finch Report, mainly because - by choosing gold – Finch recognised their need for a sustainable business model. Further commentary on subsequent events can be found here:

- PA response on RCUK Policy on Open Access and Supporting Guidance (20 March 2013)
- PA response to BIS Select Committee Open Access (February, 2013)
- HEFCE Letter: Open Access and Submissions to the REF post-2014 response from PA

5) Consultation Methodology

5.1 Approach and methods

1) A qualitative approach was taken, using focus groups and interviews with a semi-structured interview schedule. A very pragmatic approach was taken to recruitment and timetabling, accommodating the needs and priorities of participating staff.

5.2 Sample

Interviewees were drawn from a range of relevant staff across each institution. Potential participants were identified and approached within each institution to represent a range of OA stakeholders. The sample – across three focus groups and five interviews– consisted of 29 university staff. Participants were predominantly research staff from PhD to Professorial level, a mix of both STEM and non-STEM, and also included lecturers and research support. Interview participants were senior academics (research and strategic leads). *NB: Specific demographic profiles of focus groups and interview participants are not given due to the risk of identification. Quotes are similarly not attributed to individuals / data collection sessions. Analysis suggested that meaningful comparison between roles / disciplines (etc) was not viable (small sample size, mixed academic/support roles)* and that quotes were sufficient to provide the breadth of findings needed.

5.3 Ethics and anonymity

Ethics approval was granted by Coventry University Health and Life Sciences Ethics Committee. All participants were given an information sheet and consent form and given the chance to withdraw at any time. Sessions were digitally recorded and transcribed by the interviewer/focus group facilitator in each case. Transcripts were redacted as necessary to protect participant identity. To preserve anonymity in this report, quotes are not attributed to individuals or academic/non academic role. As the aim is to identify a broad range of needs rather than comparatively assess needs of stakeholder communities, anonymised reporting is both appropriate and complies with ethical approval.

5.4 Analysis

Transcripts were thematically analysed, with line by line by line coding followed by higher order categorisation. As the aim was to identify OA needs and associated solutions, findings (themes) were rooted in the data (i.e. based directly on transcripts rather than abstracting more conceptually). This was underpinned by a positivist, realist ontological approach (the stance that there *is* a truth to be known and it can be uncovered). Thus findings are presented within the interview schedule categories, with sub themes/categories itemised within.

6) Findings

Results are presented in the following sections:

1. Knowledge of OA
2. Drivers
3. Barriers
4. Facilitators and support

Where appropriate, diagrams are provided to show the links between themes / sub themes.

6.1 Knowledge of OA

Knowledge about OA processes fell into three domains:

- **'Fact':** Knowledge that was fully detailed, understood and (where appropriate) translated into practical approaches
- **'Unsure'**: Information and understanding which is assumed, arguable, opinion or conjecture.
- **'Confused'**: Misunderstood or conflicting knowledge

a) 'Fact'

Some respondents had an excellent working knowledge of OA processes, understanding the alternative paths (Gold and Green) and the associated archiving.

> *…if you have a paper accepted into a journal that is fully Open Access, they're publishing and it's completely open, that's your gold route. If you have a paper that's accepted into a journal that isn't an Open Access or you don't want to pay the additional fee for it, you can then put it into a repository ...and aside from potential copyright embargo, so it may be that it's not available for six months or whatever, it then becomes available as a green route, but the version that goes up is the final, accepted version for the journal but not once it's been formatted into that journal's kind of standard template, so it's almost like the final Word or PDFd version of it (FG2, 78-88).*

Others had a more nuanced understanding of the financial aspect, and the availability of resources to fund APCs.

> *… legitimate journals charge for Open Access and they have various different types of embargos and stuff…if you want it to be free immediately you'll have to pay quite a substantial amount of money. And a lot of national, international funding bodies will put that into the money that they're giving you, knowing that when you put into a reputable journal, they want it to be Open Access and they'll give you that money (FG2 189-196)*

One respondent outlined a business-focused solution to planning funds for OA, capitalising on funding criteria to support publishing costs:

> *[I've started] putting a clause into our proposals that says unless you allow us to put this in the public domain, we have to charge you VAT. We have costed for dissemination that means that we don't charge you VAT. If you want us to revise that, then the costs will be changed accordingly (FG1 212-220)*

However, there is also recognition that such opportunities are proportional to the size of the project and funding available:

> *But on the smaller projects, obviously you can't go for the APC route, because that would be more than 20% of the value of the project (FG1 236-7)*

This level of detail however was restricted largely to a small number of participants, with some holding considerable, operational knowledge with most demonstrating a more vague understanding.

b) 'Unsure;

Data also showed a level of debate and assumptive statements around OA paths, leading to a lack of certainty. Even where the difference between Gold and Green is understood, the merits of each path fall under a further level of scrutiny. Gold is consistently cited as the 'better' route, but with caveats over the cost that this incurs:

> *I know mostly about what the general philosophy is behind it; why the Councils are going for open access and why it's a good thing, but actual routes I'm not so sure of what we're doing. I know publishers now offer an open access option and you have to pay more for that, so that's one route and I understand the University are putting in some routes themselves with author copies (I1 22-26)*

Perhaps more insightfully, participants made reference to the relative merits of each route and the intrinsic value of each approach. The merits of Green were continually positioned only in relation to the high costs of Gold; i.e. that Green solves the financial problem, it is not a valuable route otherwise.

> *Well, the debates I've been reading about it on various kind of forums, it seems that gold is considered to be the kind of, the gold standard in terms of what you're actually giving as Open Access, is the final formatted version, as it would have gone into the journal, so it's the journal's own Open Access. As opposed to the green route being almost like the way to get around the implications of an Open Access journal, particularly in terms of financial costs (FG2 63-68)*

Lack of clarity extended to understanding on OA compliance requirements:

I think someone was telling me last week, from 2016 it's going to be mandatory? (I5 16-7)

c) 'Confused'

Inherent in much of the discussion on OA however was a sense of confusion and a lack of clarity on what OA was or how it operated.

I remain somewhat bemused by the issues involved (I4 62)

As a result discussions easily veered in allied directions such as OA publishing, data, article availability and subscriptions. In one focus group discussion a participant highlighted the divergence of conversation and the clarity on where free access was positioned:

No, we're talking about two different things, she's talking about, you know, publication on those journals, that shouldn't be free, but then access to those journals, yes(FG2,199-201)

The OA agenda was also conflated with general visibility for outputs, with online presence cited as part of overall funding compliance but with an unclear connection to OA compliance itself:

I think as a requirement, some of the, you know, organisations expect you to say, have a website for the project you're doing and that you publish your research outputs there. So it's perhaps a bit different to the Open Access of this, to the, you know, the journals, it's different, but still…(FG2 292-5)

Overall knowledge demonstrated within these groups reflects a range of understanding, from highly detailed and operational through to general awareness with a lack of clarity over the processes.

Learning for OA Workflow development (Knowledge):
The OA knowledge base varies substantially. Learning objectives for strategies to improve OA understanding in the academic and research support communities should include: 1) Understand the formal OA agenda and the definition of 'compliance' 2) Understand the Green and Gold paths with associated archiving and versioning controls 3) Understand the cost, profile and impact implications of each route

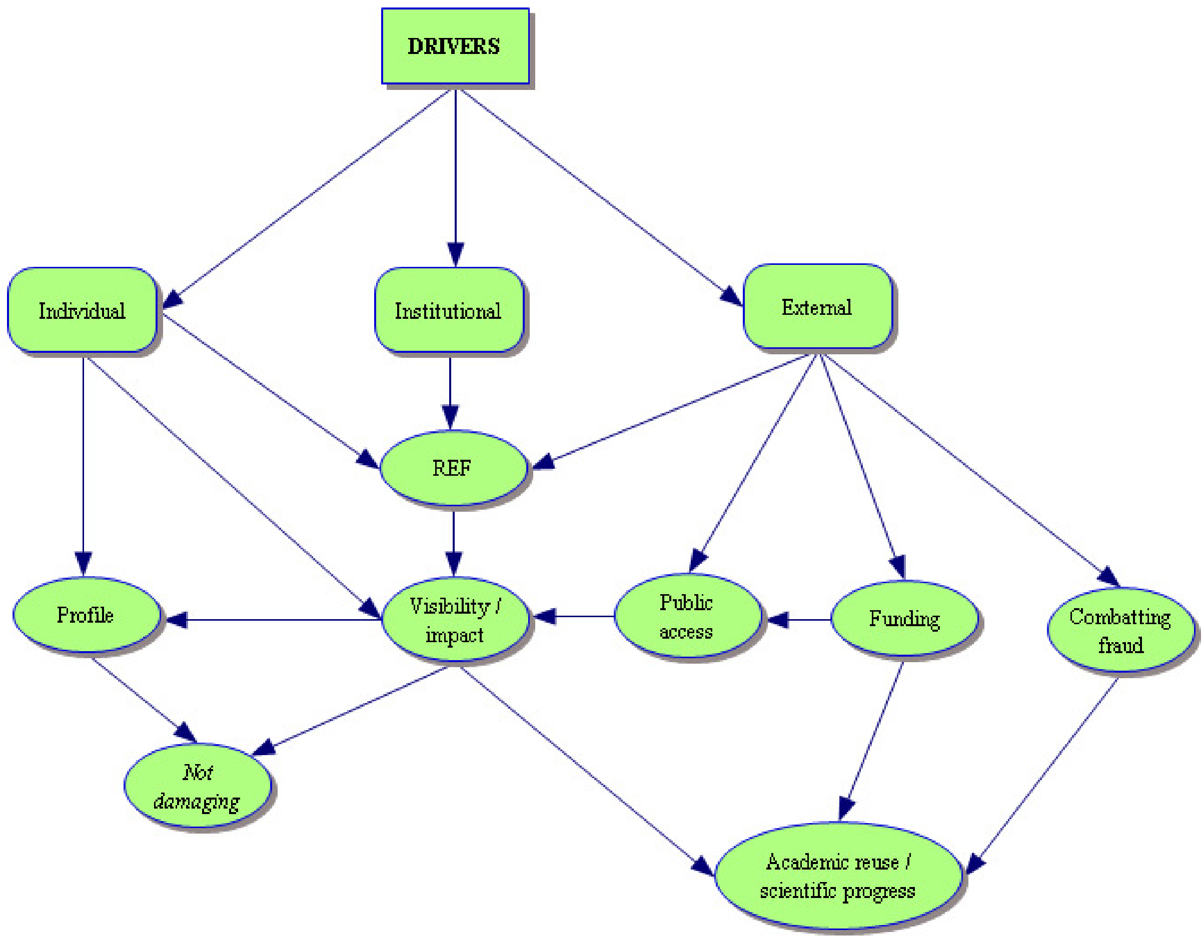
6.2 Drivers

Analysis of responses on 'Drivers for OA' revealed a three tiered but interlinked set of drivers:

1) External drivers
2) Institutional drivers
3) Individual motivations

Figure 2 shows the links between these drivers and the associated subthemes. NB: Categories are not mutually exclusive but rather indicate the root of the motivation/drive.

Figure 2: Drivers of the OA agenda and OA compliance



a) External drivers

Participants cited multiple external drivers for OA, with the main two being **REF** and **funding** mandates. REF was the most swiftly cited reason in both groups, with clear understanding that OA was fundamental criteria for inclusion in the next REF. This binary ‘in or out’ function of OA formed the basis of the drive to support OA compliance immediately:

> *If you can’t submit it to REF if it’s not Open Access, then there’s not much point (FG1 62)*

Funders’ mandates for publication and data release were understood in terms of justifying public funding and enabling **public access**.

> *… the public, in the end, are paying for, whatever route they come through … And the idea is that, therefore, the public should be able to access what they’re paying for. And I totally agree with that. I’m not sure that, necessarily, they want to, they know they can, or anything else. But they should be able to, if they want to. And I think that’s, for me, that is impactful and purposeful research. (FG1 557-64)*

> *…it’s public money, it’s Government money coming from RCUK, charity money then you’ve got a responsibility to be open and transparent. (I3, 42-43)*

This mirrored a more general view that OA supported the impact agenda:

> *…it’s really partly to promote impact, which is part of the new agenda. So it’s getting it out there and it’s optimising the readership in some way (I2 7-9)*

In parallel funders’ were believed to be driven to push secondary data / **data reuse** to both accelerate academic research and reduce duplication of funding for existing work:

> *I think the funders and, sort of, Research Councils particularly are the main driver for that, because ESRC, a few years ago, made a big thing about using more* ***secondary data*** *in their research proposals. And I think that is valid, they want more, and better use of research that they’ve already paid for, but I think it is primarily the funders that are driving that. I haven’t, I’ve heard less about making data Open Access from HEFCE, and, I don’t think I’ve heard anything from HEFCE about making data Open Access, but the big funders certainly want more [inaudible], essentially. (FG1 852-9)*

> *…if a government body has paid for somebody to do something and the data is in existence then to pay somebody else to do something either very similar, then the government quite rightly or the big funding bodies, whoever, could quite rightly say, “Actually the data’s there, if you want to look at it in a different way, fine.” (FG2 791-6)*

> *[OA] speeds up the* ***rate of progress*** *because you’ve got more people that can just jump on things (FG2 849-51)*

Academic reuse for data was also cited as a means to combat fraudulent reporting of research.

… transparency and to avoid a, kind of a misconduct I suppose. There have been a couple of cases recently where papers have been withdrawn from quite prestigious journals because it's turned out that there's been falsification of data (FG2, 778-83)

There were however concerns that the availability of data may simultaneously support funders to finance more original research, but may reduce funding available for similar work **(duplication)** or replication:

And also, you were talking about the funding bodies, if they, they don't want to be duplicating stuff that's already been done (FG2942-6)

You have to demonstrate when you are putting a research council application in that the data you are seeking is new and original and it's not already out there (I5 173-5)

These external drivers dovetailed with institutional and individual drivers most prominently via the **impact** agenda and perceived benefits of **research visibility**:

...the greater good requirement is to be for others to be able to see the research others have done and build on it. Because the more open we can be about our results and our research the better it generally is for the scientific effort and so on, so that we can get really good research done as a country, and world really because when they're open of course, they are open to everybody (I1 8-12)

We might be looking at a REF point of view at impact outside of academic impact, but there is still having academic impact….being able to see the research that comes off your initial research is still a good thing (I3 90-93)

b) Institutional drivers

With **REF** as a primary external driver, this was also positioned at the institutional level with respondents citing the need to facilitate their institutional submission as a primary driver:

… the main reason that we want to do Open Access is to get your work out there but ultimately, like I said, the next REF, I'm looking at the perspective that we need people from our organisation in the next REF submission and they're going to put restrictions on that only things that have been made Open Access will be admissible. (FG2 216-221)

Responses highlighted that for some, the strategic aims of the institution overrode any personal or wider, external agendas. More simply, institutional decree was the sole driver of OA compliance for some individuals:

I do it because it's what we're asked to do as an institution, rather than for any personal reason really (FG1 501-2)

Because it underpins the aims of the institute. You know, the institute is aiming to improve the health and wellbeing of the population (FG1 518-9).

c) Individual drivers

Alongside the wish to be included in the REF, individuals were mostly motivated by the contribution of OA to profile building and enhanced visibility of research. Multiple co-related drivers were presented, covering an overall belief that OA enhances / enables public access and the speed of dissemination:

I mean you want your publication to reach, I mean, as wide an audience as possible, so Open Access is that way (FG2 105-6)

I've worked with various community stakeholders, and they want access to the research. And they always complain, like, oh, we can only get reports, we can't get access to the peer review publication, because we have to pay and we don't have money to pay for that. So I think that is, I agree that it should be made publicly available (FG1 551-5)

Publishing in Open Access journals is…much quicker, so your publication gets out their much faster (FG3 587-8)

In parallel this visibility and access may stimulate interest by funders and attract future research opportunities:

within the local communities…if there is Open Access to the work that we do with the university, and its reputation, commissioners might come more readily to us to have research done here, when they see that we are producing research with value (FG1 75-9)

if you're doing a sort of programme of research, if you need to demonstrate what you've achieved… then you will go for Open Access, if you can, because…you'll get them out there quicker and then you can reference them in your funding applications for the next stage (FG3 599-603)

Respondents also noted the potential within-academia benefits, with higher visibility through OA potentially generating more citations.

And there's more chance of people noticing you because it becomes a public asset, freely available to everyone. More people are likely to have a look at your research and you're more likely to be cited than if you had suddenly taken your work to subscribe only journals (FG2 32-5)

…the likelihood that people will read it and cite it and it'll influence more people's work (FG2 98-9)

The more accessible your research is, the more you enhance your profile. (I4 31-32)

Ultimately the individual drives reflected the implicit link between OA and impact:

> *If it can't be accessed in the Open Access arena, then it, then that limits the impact that it can have (FG1 71-2)*

Learning for OA Workflow development (Drivers):
There are multiple and interlinked drivers for OA within the HE community, with REF and funding most strongly persuasive. Operational objectives should therefore include: 1) Clarify and emphasise the REF requirements for OA 2) Clarify and emphasise specific funders' requirements for OA 3) Clarify and emphasise the benefit of increasing access to articles via OA routes (including visibility, speed of dissemination and citations, impact and positive public reception) 4) Communicate a clear and consistent institutional message on the value of OA

6.3 Barriers

Participants highlighted a large variety of barriers to OA publishing and data release. These are summarised in Figure 3 overleaf. Factors are often interlinked, and whilst linearly presented reflect a larger and complex set of variables underpinning OA activity.

a) Money

Funding for OA is seen as a complex issue, comprising a mix of practical limitations, funder mandates and broader attitudes and concerns. Fundamental issues on the availability of funds for Gold were:

The overall **limited funds** available, along with concerns over where funding could be found:

> *[funders] might require it but the pot is only so big (FG2, 389)*

> *I guess our main concern would be the cash, at the end of the day…if we want to publish Open Access, we have to spend pretty much our entire research budget for [redacted] on publishing one article (FG 1 335-9) .*

Specific issues relating to funding from **charities**:

> *charities who are funding research aren't, they don't really mind that much if it's Open Access, they don't stipulate it so in turn you can't then go ask for a couple of thousand to do it (FG2 403-5)*

And a broader concern about those **researchers who do not have access to any relevant funds:**

> *if you're between jobs or that sort of thing, or relating to a PhD and wanting to publish, again getting your hands on the money to do it in an Open Access way is a real pickle for these sorts of, people in these sorts of positions. And you know, this is really, really detrimental to the people wanting to start off getting an academic career (FG2 422-6)*

Financial concerns were cited frequently and early in discussions, suggesting their salience in attitude formation about OA.

b) Planning

Inextricably linked to financial considerations is the ability to plan OA publications. Aside from the mechanisms to fund OA outputs and the pro/anti-OA attitudes to decisions, the process of planning is difficult. Much of the challenge relates to the **difficulty in establishing a likely cost** of publication from the outset, with some describing this as a *'finger in the air, an average amount'* (FG1 117-8). However, the need to do so was understood:

> *you've got to second-guess it up front, it's no good going in at the end and saying, "Oh, by the way..." (FG1 111-2)*

A related concern revolved around the **changeable and unexpected** nature of research and the how planning may not match the ultimate journal submitted to.

> *it might not be the journal that you end up publishing in (FG1 143-4).*

> *Cause one of the problems…is that even at the outset of the research, when you're putting your proposal in, you might say, "Oh, well, I want to publish an article in the journal of," and you can look up how much they charge for gold standard. But when it comes down to it, when the research is done, and when you're actually writing the articles, you might go to a totally different journal, and that might have a different pricing structure (FG 123-9)*

> *it's going to influence which type of journals we're going to be able to put papers into because some funding you do bid for with something built in to pay for those but if you haven't already kind of pre-empted that it becomes unaffordable (FG2 26-9)*

OA discourse is to a large extent underpinned by the belief that there is a finite 'pot' of money available, easily exhausted by Gold routes and can counterproductively limit outputs as a result:

> *I think one of the concerns of Open Access, it would limit the amount of publications someone can do within a certain period of time especially if their project is, they are investigating a very fertile topic that can generate many papers, they'd be limited by the budget, publication budget, of how much they can publish (FG3 380-4)*

The overarching dialogue about planning financially related to the explicit cost of Gold rather than the lower (free) cost for Green. Arguably therefore the financial barriers to Gold OA mask the opportunities afforded by Green, and lead to a higher perceived complexity in planning OA outputs as a result.

c) Attitudes

Participants' attitudes to OA funding and the associated planning showed concerns over both the **impact of building costs** into bids...

> *bidding is already difficult enough and people look at, you know, what you're putting in as opposed to what other people are putting in and if you put this additional cost in (FG2 378-80)*

...alongside the **strategic decision making** which occurs as a result of building in a paywall on the academic (rather than reader) side and the dilemma between **specialised vs high impact factor** outlet:

There's one specialist journal I know that I would like to publish in …. I sort of raise that at my [name of group and [will be told] "Well there's no money to publish there," but it's quite a good journal, all those kind of specialists in my field…and it would be useful place to put an article to raise my profile. But … it's not seen as high enough credibility journal to warrant, you know, the university…spending it on that journal (FG2 355-65)

More fundamentally, attitudes to OA reflect a perceived stigma of green publishing by the academic community:

*the **academic snobbery** aspect, that if you didn't publish it the proper way, the gold way, that perhaps it could end up with less credibility because it's gone through the green route (FG2 478-80)*

d) Output

The nature of the output(s) adds considerable complexity to OA publishing, automatically producing a series of points at which barriers occur. Traditionally journal articles were a single item, with the journal copy the only version and physically available only at the point of publication. However, with the advent of electronic journals and the OA agenda, versions exist pre-print, may not be physically printed and require a version to be deposited in an institutional repository. The situation is complicated by the publishing transaction occurring between the author and publisher, aside from university input. The upshot of this is that increased and unprecedented **responsibility is placed on the researcher** to comply with guidelines and the risk of having unintentionally *'transgressed' (FG1 393).*

A primary source of discomfort was the **credibility of OA journals**. With the emergence of OA, authors have become increasingly concerned about the rise of fraudulent journals and the difficulty determining which are legitimate. This difficulty is compounded by the lack of 'quality' markers and self-branding:

a lot of websites, they claim they are digital libraries or whatever (FG2 147)

there are a few dodgy journals, they…try to make money from…publishing your papers (FG2 137-9)

the sudden push for Open Access means there's a whole load of new journals that have appeared and trying to kind of navigate your way through the credibility of them because especially if they're new journals, you can't necessarily rely on looking at an impact factor to see how credible they are because you don't expect that to really be effective for a few years until the citations have started, so knowing whether to put your eggs in that basket or that other one is not going to be easy (FG2 727-34).

Such concerns complicate an already difficult process for choosing a high quality journal:

… bottom line is getting it published. If I can get it published Open Access, then that's a boon. But, you know, if it's either Open Access or nothing, then that's going to give us a real problem....because some of them…have got very specific audiences, or specific type of research or evaluation, and the appropriate outlets are rather limited (FG1 259-64)

There are also negative implications for disciplines and sub-specialities resulting from the OA funding model:

….there is a major concern in the humanities and some of the social sciences, and that is that professional societies depend on their journals for a source of income and if that business model is undermine then some professional associations would have serious concerns (I4 22-6)

Open Access compliance adds a new tier of effort and complexity, with versions of articles to be **deposited in an institutional repository**. With HEFCE rules on article visibility, this is a time urgent activity and requires authors to (a) know the rules and (b) action accordingly:

the onus is on the researcher, at whatever level, to send it on acceptance, not at the point it's published (FG1 418-9)

whether there are any specific procedures that we should be following now, I'm not sure whether those are in place or not, or perhaps that's something we are currently developing at the University (I1 64-67).

In contrast to the urgency, participants' responses suggested a level of ambivalence and inaction which led to inconsistent depositing/archiving:

I've published about three or four articles … that really ought to be on there, and I know they're not, and I haven't put them on. So the titles are on, but not the actual, I haven't actually posted the article on (FG1 359-62).

I, every now and again I do a, sort of, blitz, in actually sending my references through, but then actually, there's a bit of a delay in sending PDFs (FG1 369-71).

Such responses were qualified by authors by acknowledging the **copyright** hurdles to depositing:

But my excuse for that, apart from the fact that I forget to do it, is, I'm actually not sure of the copyright (FG1 371-2)

I send the wrong one, anybody will check it? (FG1 384-5)

It's not a simple matter of, oh, here's the PDF. You've got to actually go back to the journal and check what the copyright is, and when can you send it, and …you think, oh, I can send it in two month's time (FG1 399-403)

The resulting multiple **versions** (journal version plus institutional version) also triggered concerns about harmonising **citations** and negative effects on paper **aesthetics**. With a journal and an archived copy, citations to the printed paper may be reduced /masked by citations to the archive copy:

> *particularly students who use Google Scholar a lot will be flagging up your green route work possibly and citing that and does that then get cited differently than the actual publication?(FG2 594-98)*

Equally the presentation and pagination differences between a pre-print and print version can cause issues with citations:

> *But if you're citing specific pages within then they will be different and you might get confusion about some, so-and-so saying something on page 12 in the green version, which is actually on page 15 in the other version, but actually page 12 makes perfect sense in a different context and suddenly you've created a, you know, sort of, a mis-communication (FG2 522-5).*

This combination of factors can make academics uneasy about the archiving process and detract more from Green as a viable option.

> *It feels slightly uncomfortable putting something up that I know is not quite right…that I know there are minor errors within it (FG3 288-9)*

Finally, when juxtaposed with other 'visibility routes', institutional repository deposit can seem less appealing and thus reduce likely compliance:

> *If I'm talking to people outside the university they're more likely to be on ResearchGate than they are to go looking for [repository] … nobody asks you, "Have you got an institutional repository where your papers are held?" …they'll say, "Are you on ResearchGate?" …and ResearchGate will ask me. If I have something new published I get an email saying, "Is this your publication?" - [the repository] doesn't do that so it doesn't prompt me to go into it and that's probably why I haven't visited it for a while (FG3 539-55)*

Data suggests a compliance risk generated by limited understanding of time urgency, ambivalence through lack of knowledge and attention focused on journal choice. Versioning and the multiple copies this produces also reduce acceptability of (especially green) OA

e) Research data

The potential need to release research data also concerned many of the participants. A key anxiety was **ethics**, predominantly the need for informed consent to make data available. Participants recognised the necessity for this but also the difficulties of obtaining it:

*So you can **anonymise** anything, but, you know, if you say to somebody, "This is going to be confidential," I don't think they would expect that that would be made available to anybody else, anonymised or not (FG1 839-43)*

…and the preclusive ethics of existing datasets:

And, but if it is going to be deposited somewhere in an Open Access data repository, then I think you've got to get the participants' approval. Or the, you know, acknowledgement that they are approving, in that knowledge that that's going to happen. And I don't think that would be a huge barrier, but you can't do it retrospectively (FG1 810-12)

…with qualitative the challenge of being able to give enough data for the analysis without compromising [confidentiality] (FG3 636-8)

Ethical concerns extended to the challenge OA data releases poses for drawing meaningful conclusions from situated research. Re-use can **decontextualise** data and strip away the meaning, even when only mildly reworked:

you could have somebody missing a subtle nuance into the way that you've structured your data table and completely misinterpreting the numbers that they've just taken out of context and that could have some really big kind of ramifications to it, particularly if you go into medical fields and start drawing conclusions there (FG2 932-6)

Safeguards to support the protection of data require considerable effort, and even when academics engage in the OA agenda, this workload may be prohibitive:

I'm very keen on open access for papers...I'm not so keen on publishing the data really. What are my reservations? First of all, it's a lot more work to make sure you've got that data in a format that is going to be understandable by others. You've got to be absolutely rigorous that that data doesn't contain anything that could be confidential.... And I wonder how many people would use it. (I1 140-8)

With the open data there is a lot more danger if you're not managing your data correctly, if you're not taking names out and things like that, if you don't have experts around managing it (I3 113-5)….. the expectation will be on PIs to say that you need the expertise to be able to manage your data. And that's a much bigger ask, and it's a much bigger culture change as well (I3 118-20)

I'm not a particularly statistically competent person…I'd have to be pretty damn sure I was accurate in whatever I was doing or saying (I4 95-97)

Alongside **consent** from participants and appropriate reuse, data suggested researchers were concerned for their own level of consenting for reuse. Falling under the banned of '*intellectual property*' *(FG2 881-2),* a particularly strong concern was the **loss of control over data** once released, potentially **preventing self-reuse** and allowing others to use the data as they wish.

> *You don't know how people are going to use it, for what purpose…and how they are going to interpret it (I5 195-6)*

> *And so what happens if you publish your data in one publication because you've analysed it a certain way and your intention is then to use a different type of analysis to pull out a different aspect of it…*a*nd meanwhile somebody else could have taken your data and gone and done that and then published it (FG2 817-24)*

> *…if I did generate some stuff myself, all of a sudden now I've got to worry that it's going to be taken on much faster by another group (FG2 860-2)*

> *Every university's in competition with the next university about the number of publications they have, the number of research grants they bring in and if somebody takes your work and then goes and uses it to do the things that you were about to do, then…that might close down the whole [programme]{FG3 791-4)*

Taking this further, one participant highlighted the complex landscape of OA data, flagging the '**enforced collaboration**' which is generated by data re-use:

> *I think there is an element of this ownership and I want to choose who I collaborate with rather than any old person just sort of coming up and perhaps not thinking in the same … it just worries me (FG2 973-5).*

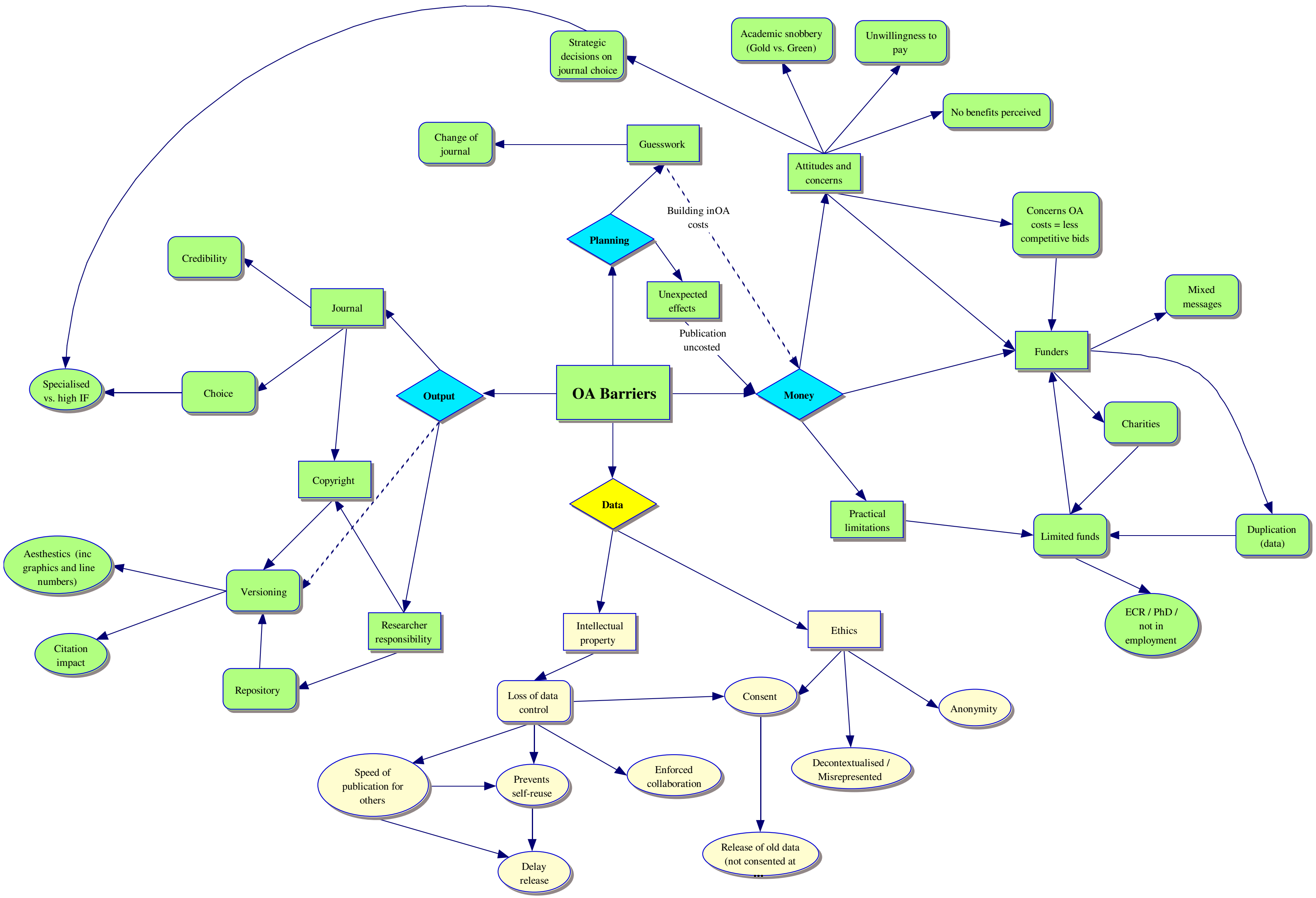
As a consequence of these concerns, participants noted that academics may use **delaying strategies** to limit the risk and release only when necessary. This protective stance ensured compliance whilst maximising academic opportunities for the researcher themselves:

> *Which in turn could slow it because what you'll end up doing is researchers will sit on their data until the very last possible minute, when they've done every form of analysis they could ever think of doing and they're ready to publish everything, so that nobody else can swoop in and take it (FG2 864-7)*

More positively respondents also noted that, if institutionally supported, ethical concerns could be managed and data released in an appropriate way:

> *So I'm saying, as long as the whole issue of the governance or management of making data available is not prescriptive, in a sense that it's regulated so you have to do it, then I think that's fine (FG2 834-6)*

Figure 3: Barriers to OA



Learning for OA Workflow Development (Barriers):

Multiple barriers exist for OA, consisting of a mix of pragmatic, attitudinal and behavioural. Much of the concern over OA relates to costs and planning for outputs which may change over the course of the research. Data release attracts a specific set of concerns, centred around consent and the loss of academic control upon making data available.

Strategies to reduce (perceived) barriers should include:

1) Reduce the assumption that 'OA is costly' by clearly articulating the differences between Gold and Green
2) Challenge the default deposit to external websites (e.g. Researchgate)
3) Clarify the funding available for OA and determine the processes for allocation
4) Develop internal strategies for delivering OA without external funding (e.g. when working with charities)
5) Upskill funding teams with OA planning strategies (e.g. how to build in costs when eligible, advising academics when funding is not available, determining restrictions based on funder guidance)
6) Increase awareness and acceptance that the Green route is a viable alternative to Gold (accompanied by clear institutional messages that Green is not substandard and Gold is 'by exception')
7) Develop as far as possible up to date lists of journal costs to support planning activities
8) Review the decision making strategies at departmental / faculty / unit level for allocation of funds (including deciding what is funded, what is not, and criteria why)
9) Develop clear technical and behavioural workflows for paper submissions; this must include guidance on copyright, process for depositing on repository (and when), and help with versioning
10) Clarify the implications of archiving a pre-print version on a repository
11) Set institutional guidelines, safeguards, support and resources for data management/release (quantitative and qualitative) and preserving anonymity/confidentiality. Guidelines may need locally translating, but overarching concerns over consent and ethics need governance.

6.4 Facilitators and support

The most frequently cited facilitator for OA was **information provision**. More specifically staff wanted information on:

- Best journals available (and emerging good journals) (e.g. FG2 1151-60)
- Costs for specific journals (to build into bids) (e.g.FG1 638-9)
- Signposting to other relevant information for informed decision making (e.g. to repository team) (e.g. FG1 681-92)

This was underpinned by a broader awareness raising and knowledge building need, the aim of which would be to encourage more institutional OA activity:

> *Make sure people know it [information] is there, because we want to promote what we have…we want people to be more aware and to be encouraged and supported to put out their research, be it in a form of publication or just the raw data aspects of it, we want to know a little bit more so we can perhaps contribute towards the, achieving the aims of the university (FG2 1211-6)*

This information must be accompanied by **clear institutional strategy** and **unambiguous messages about OA compliance**. Moreover this message needs to be weaved throughout the organisation to ensure a coordinated approach:

> *The message needs to be consistent right across the board, not just to researchers, but also to the bidding office, to the finance people (FG1 694-5)*

Discussions evolved further into mechanisms by which information could be most fruitfully shared to support this coordinated approach:

> *And I think that's really important, that the support organisations, the support mechanisms within the university are aware of that, and the implications for each of us. But also it's about sharing information … policy or good practice that could then become policy across the institution …it does need that, sort of, recognition, at an institutional level (FG1 708-21)*

Participants also felt that such information needed accompanying with clear **guidance** and **workflows** to support academic to comply most simply. This ranged from basic techniques such as having text available on the repository for inclusion into bids (*e.g. FG1 967-7*) to more technical and behavioural procedures:

> *…as soon as it's been accepted in its final form… send that to the repository, and have an embargo at the repository level…You could then have something [on the repository] that's, sort of, kept on embargo six months, or two months, or whatever (FG1 405-10)*

> *we do need some sort of slick system I think in place to remind everybody or to make sure we get them deposited (I1 42-43)*

In parallel there is need for training and support for researchers to facilitate not just mandate OA compliance:

> *there does need to be a level of consistency of support, as well as just expecting the researchers to be doing it (FG1 705-6)*

As responsibility lies with the researcher to instigate the process of OA (alerting/depositing from acceptance), the success of approaches rests dually on ensuring academics behave consistently and compliantly:

> *…proactivity required on the part of the researcher, isn't it, to actually remember, and to do the placing of the article on the repository (FG1 348-50)*

Alongside institutional commitment to supporting OA structures:

> *It's quite clear that the University needs to resource this area effectively, full stop.(I4 110-1)*

Learning for OA Workflow Development (Facilitators and support):
Support for OA is primarily delivered by overcoming a range of barriers and providing information and procedural clarity. In addition to those approaches listed under 'Barriers', facilitative strategies should include: 1) Provide clear and unambiguous information on OA, including costs and support opportunities. 2) Improve overall engagement with the OA agenda and increase the willingness to comply 3) Produce and communicate clear workflows (technical and behavioural) to simplify the perceived process and reduce errors

7) Conclusions

Whilst there is considerable activity in the OA domain through publishers, HEFCE and funders, there remains concern in the academic community. Anxieties relating to OA paths and release mechanisms hinder engagement, although the overall benefit of OA is - to a large extent - understood and accepted. Rhetoric about impact, visibility, public access and accelerating scientific development underpin the willingness to engage, but pragmatic issues (costs and procedures) and ethical concerns prove a barrier to action.

As OA compliance is a complex area, requiring a mix of behavioural and technical actions, strategies to increase adherence must be equally multifactorial. Approaches must include:

- Clear and unambiguous institutional strategies, communicated across the organisation and consistently implemented.
- Up to date and available information on credible journals, costs and practical issues (e.g. embargo periods)
- Training, guidance, workflows and procedures to support academics navigate OA requirements
- Safeguards for data protection and ownership, and associated incorporation of ethics concerns at the planning stage (consent)
- Challenging assumptions over the credibility of the green route and increase acceptance
- Where funding is available, the allocation of money must be transparent and consistent
- Where funding is not available, the optimal strategy for publication via green must be clarified and communicated
- Reduce unnecessary effort (e.g. duplication) and streamline processes for copyright confirmation, repository deposit and embargo compliance.

It is notable that participants did not cite specific technical solutions/workflows; arguably this reflects a lack of insight into the process to be automated and an area in which repository officers and system specialists are invaluable. An aerial view, supported by a range of stakeholders – users, developers and library experts alike – is needed to develop the vital coordinated embedded approach.

8) Next steps

The learning from this report will form the basis for workflow development in WP2. Needs will be mapped against existing processes to generate a change strategy and address the key areas of difficulty. In parallel, learning will be shared across the Pathfinder community to support collaborative developments.