



**UCL**

## **Influencing the Deposit of Electronic Theses in UK HE.**

**Report on a sector-wide survey into thesis deposit and open access**

Josh Brown (SHERPA-LEAP) & Kathy Sadler (UCL)

October 2010

**JISC**

# Influencing the Deposit of Electronic Theses in UK HE.

## Table of Contents

- 1. Summary.....3
- 2. Introduction .....5
- 3. Methodology .....5
- 4. Participation .....6
  - 4.1 Respondents by Group .....6
  - 4.2 Respondents by number of doctoral students .....6
- 5. Acquisition of e-theses.....7
  - 5.1 Deposit.....7
  - 5.2 Digitisation.....8
- 6. Open Access.....8
  - 6.1 Formal policies and open access.....8
  - 6.2 Making theses available.....8
  - 6.3 Repositories and e-theses storage .....9
  - 6.4 Restricting access to theses .....10
- 7. Policies and Practices.....11
  - 7.1 Formal policies .....11
  - 7.2 Workflow analysis .....11
  - 7.3 Collaboration.....12
  - 7.4 Metrics.....12
- 8. Barriers to e-theses deposit.....13
  - 8.1 Impact on Future Publication .....14
  - 8.2 Third Party Copyright .....16
  - 8.3 Plagiarism.....17
  - 8.4 Sensitive Content.....19
  - 8.5 Practical barriers .....21
  - 8.6 Other concerns.....21
- 9. Conclusion .....22
- 10. Recommendations.....23

Version 1.0  
October 2010

# Influencing the Deposit of Electronic Theses in UK HE.

## 1. Summary

### Background

This survey formed part of the 'Influencing the Deposit of Electronic Theses in UK HE' project<sup>1</sup>, commissioned by the JISC and led by UCL. The survey was designed to capture a snapshot of current and planned electronic theses policies and practices in UK HEIs, and to gather evidence about the main barriers to the electronic deposit of e-theses.

The survey ran during June 2010. The 155 UK HEIs that award or plan to award research degrees were targeted. 144 responses were received, a response rate of 93%. Respondents included 19 members of the Russell Group, 17 members of the 1994 Group, 23 members of Million+ and 21 members of the University Alliance.

### Main findings

#### *Deposit of electronic theses*

Currently, 90 responding HEIs (63%) accept electronic copies of research theses, 70 through formal policies or mandates. A further 35 HEIs report plans to introduce the deposit of electronic research theses within a five-year timeframe. Of the group of 19 HEIs with no plans to begin collecting electronic theses, the highest number of doctoral degrees awarded was reportedly 10 per year.

Within five years, therefore, electronic thesis deposit will be accepted in at least 125 of 155, or 81% of research degree-awarding HEIs. Using HESA statistics<sup>2</sup> this indicates that, within five years, copies of over 99% of the UK's annual output of PhD theses will be held electronically by awarding institutions.

#### *Access to electronic theses*

62 out of the 90 HEIs (69%) that currently accept e-theses make them openly accessible, either through a repository (56) or a website (6). In the remaining institutions, e-theses are stored offline or are subject to permanent access restrictions.

90% of the formal policies or mandates reported are 'opt-out': electronic theses are placed on open access unless a student requests an exception.

#### *Barriers and concerns*

Respondents were asked about levels of concern within their institutions about the most frequently raised issues relating to e-theses. The survey found high levels of reported concern about the risks associated with third party copyright infringement in electronic theses (89 HEIs); about plagiarism (76 HEIs); about the inclusion of sensitive data within theses (75 HEIs); and that open e-thesis deposit might hinder an author's future publication prospects (72 HEIs). However, very few instances of copyright-driven 'takedown', plagiarism, issues arising from data sensitivity, or publication refusal were reported.

#### *Other issues*

The survey found that fewer than 30% of HEIs with e-thesis collections maintain usage statistics. There is inconsistency of accounting method among those who do, and not all common repository platforms readily support the collation of access statistics.

Other reported practical barriers to the systematic collection of electronic theses were resource constraints, unsurprisingly, and a lack of institutional priority or policy delivery. Of the few other

<sup>1</sup> Project website: <http://www.ucl.ac.uk/ls/etheses>

<sup>2</sup> Figures based on HESA statistics for qualifications obtained in academic year 2008/9. Data table available online at: [http://www.hesa.ac.uk/index.php/component/option.com\\_datatables/Itemid,121/task.show\\_category/catdex,3/#quals](http://www.hesa.ac.uk/index.php/component/option.com_datatables/Itemid,121/task.show_category/catdex,3/#quals)

concerns that were reported, the challenge of streamlining legacy procedures for the electronic environment was the most commonly expressed.

## Conclusions

The electronic deposit of theses, actual or planned, was found to be commonplace within UK HE, and is likely to embrace nearly all the sector's annual doctoral thesis output within five years. There is little need to attempt further to influence HEI decision-makers to accept the deposit of electronic theses.

Levels of concern about the most frequently identified 'barriers' to collecting and exposing electronic theses are high. The survey also showed that these barriers are overstated in some cases and manageable in all, through proven strategies involving advocacy, graduate education and measured use of temporary embargo.

The electronic corpus is not yet fully accessible. Among the HEIs that have adopted formal policies, 'opt-out' open access is usual. The wider adoption of formal policies with 'opt-out' open access should be encouraged. The survey shows that the perceived barriers to such policy-making are evidently tractable. Sharing successful workflows, good practice in dealing with the common concerns, and proven and mature OA-friendly policies would help to build confidence and ease and harmonise the management of e-theses across the sector, paving the way to the wider adoption of formal policies that would ultimately boost the open availability of the UK's doctoral output.

Issues of impact measurement, digital preservation and the management of multimedia degree submissions also give rise to some recommendations.

## Recommendations

### *Measures to increase open access to UK e-theses*

1. HEIs should be encouraged to adopt formal, OA-friendly 'opt-out' policies. To help to achieve this, work should be carried out to help HEIs to share experience and build confidence in dealing effectively with the most common concerns around e-thesis management, facilitating the sharing of good practice, successful workflows, and mature and proven policies.<sup>3</sup>

2. The project found no hard evidence that open access thesis deposit prejudices an author's prospects of publication. Follow-up work to investigate publisher attitudes could be undertaken; the creation of a central knowledge base (or the extension of SHERPA/RoMEO) might be appropriate.

### *Impact*

3. The ability to demonstrate the impact of open access theses is potentially beneficial to institutions. Few HEIs maintain relevant statistics; there is inconsistency of method among those that do. Standard metrics for e-thesis access should be agreed; repository plug-ins for the most common platforms developed; and guidance on the effective use of third-party resources such as Google Analytics and Google Webmaster Tools should be collated and shared.

### *Looking ahead*

4. HEIs are beginning to move away from archival thesis storage in print form: 8 HEIs already no longer accept theses in hard copy. It would be timely to explore strategies for helping HEIs to meet the challenges of the long-term digital curation of electronic thesis collections.

5. The quantity of non-text theses, such as multimedia submissions for performing arts degrees, and supplements to text such as datasets, simulations, etc., is likely to increase as e-deposit rates grow and the availability of enabling technology continues to expand. Investment in co-ordinated 'best practice' resources to support the sector in meeting the challenges of managing non-text theses – issues such as format selection, metadata standards, preservation strategies, and so on – should be considered.

---

<sup>3</sup> Certain planned outputs from this project will assist with the delivery of this recommendation: a suite of evidence based best practice summaries on the common concerns; a suite of case studies; targeted enrichment of the EThOS toolkit; and an RSP event scheduled for 2011.

## 2. Introduction

In 2010 the JISC commissioned UCL to conduct a landscape survey of thesis deposit. The intention was to gauge the level of uptake of open access e-theses policies within the UK HE sector, and help to identify and overcome the perceived barriers to the open access repository deposit of e-theses at UK HEIs.

The survey forms part of the 'Influencing the Deposit of Electronic Theses in UK HE' project, whose aim is to encourage HEIs to prioritise changes to their policies and practices that result in electronic versions of their new PhD theses being available on an open access basis from their institutional repositories. If open access is not possible, then at least a bibliographic record should be available in most cases. The UCL project is one of a pair of concurrent and complementary e-theses initiatives, the other strand being the British Library's 'Gathering evidence of the benefits of increased visibility and impact of open access theses' project, concerning the impact of the EThOS service.

## 3. Methodology

An online survey was designed using SurveyMonkey containing the following sections:

1. Policies
2. Practices
3. Barriers to e-thesis deposit

The survey contained 65 questions overall, with built-in logic to present respondents with targeted questions based on previous answers. For example, in order to identify practices and issues specific to e-theses, those accepting print deposit only were often questioned separately from those with some form of electronic deposit.

The survey was piloted by the SHERPA-LEAP E-theses working group and members of the Project Board. The survey went live on 27 May 2010 and ran until 22 June 2010.

### Scope

The target group comprised all UK HEIs who offer postgraduate research degrees. In order to achieve a high response rate, invitations were sent to one named individual at each institution. Of 169 universities and HEIs in the UK there were:

- 14 institutions identified as not offering doctorates: these were excluded
- 105 institutions where a repository manager/officer/contact could be identified
- 50 institutions where the SCOUNL representative was approached

155 invitations were emailed, one to each target HEI. Progress was monitored and up to two personal chasers were sent to maximise responses.

### Results

The key results are presented in this document; full responses are available on the project website<sup>4</sup>. Results have been anonymised and respondents are identifiable only by their university group. The primary focus of the survey was to investigate electronic theses, and findings concerning print theses are reported here only for the purposes of comparison.

---

<sup>4</sup> Project website <http://www.ucl.ac.uk/ls/etheses>

## 4. Participation

144 responses were received from the target group of 155, giving a 93% response rate. 86% of respondents completed the entire survey.

### 4.1 Respondents by Group

Respondents were asked if their institution is affiliated to a university grouping. The main university groups were all well represented with 19 Russell Group respondents, 17 from the 1994 Group, 23 Million+ members and 21 from the University Alliance.

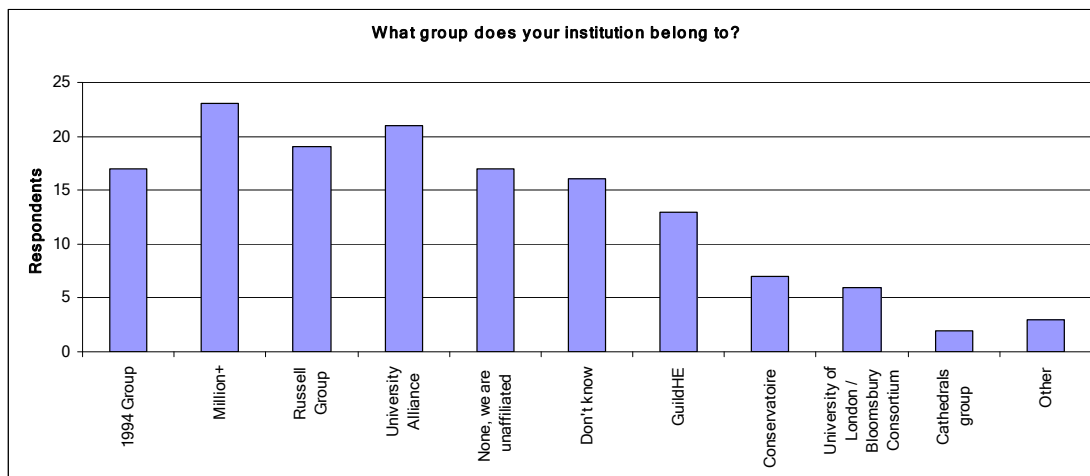


Figure 1: Respondents by university grouping

### 4.2 Respondents by number of doctoral students

Respondents were asked to identify the number of postgraduate research students in their institution. 70% have fewer than 1000.

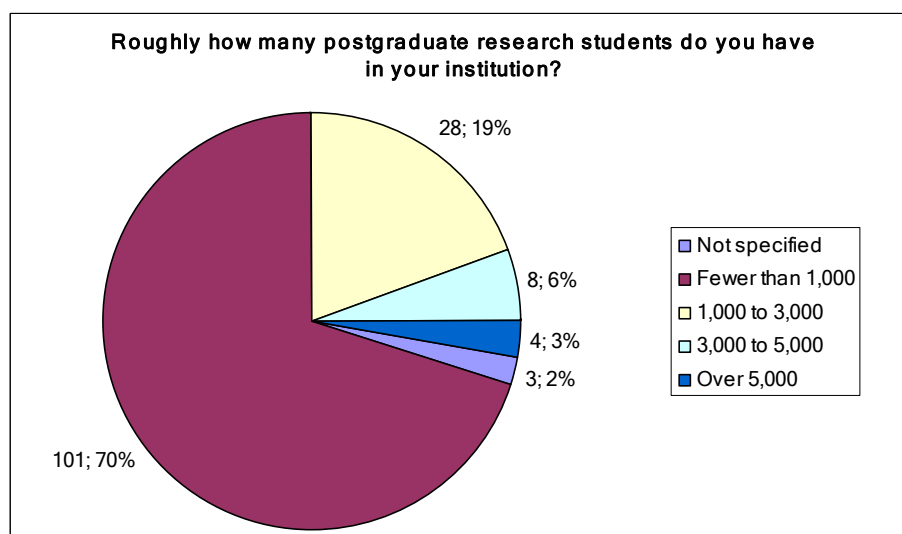


Figure 2: Respondents by number of doctoral students

## 5. Acquisition of e-theses

### 5.1 Deposit

63% of HEIs already accept some form of electronic deposit. 8 HEIs (6%) require electronic deposit only and do not accept print copies.

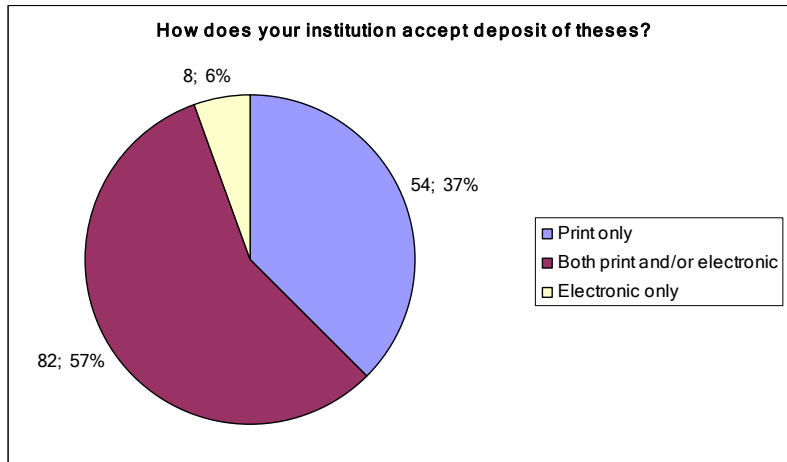


Figure 3: Method of thesis deposit in 2010

Of the 90 respondents accepting electronic deposit, 70 do so via a formal policy (mandate). The remaining HEIs have an informal or voluntary arrangement.

Of those with a formal policy of electronic deposit, 65 HEIs (93%) implement institution-wide policies.

Respondents not already accepting electronic deposit were asked about their future plans. 35 institutions plan to implement electronic deposit within a five-year timeframe. Taking current and intended e-theses deposit together, at least 81% of all higher degree-awarding HEIs (87% of responding HEIs) will be accepting the deposit of e-theses within five years (i.e. before 2016).

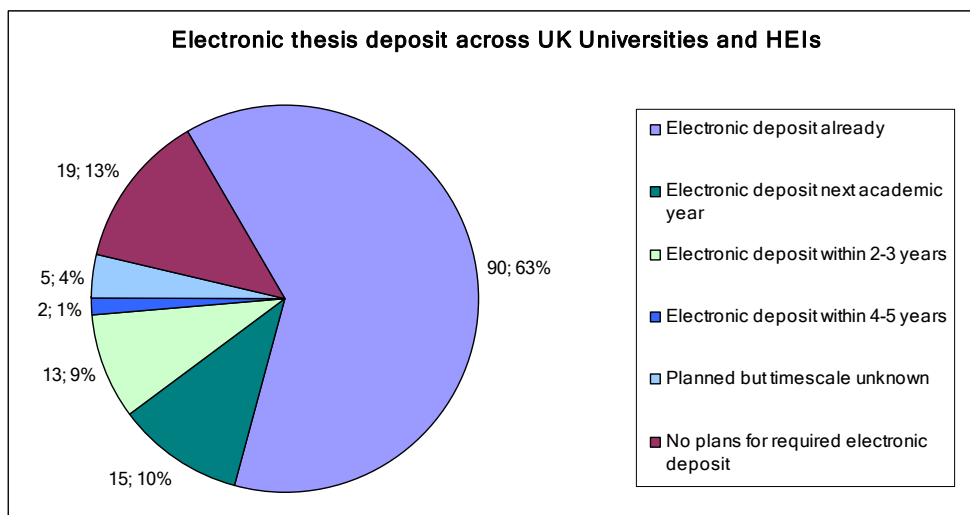


Figure 4: Electronic thesis deposit across UK HEIs, 2010 and future plans

19 respondents have print-only deposit policies and no plans to adopt e-theses. Those who gave permission were contacted for further information about their reasons for continuing to receive theses in print format only. A typical response was:

*"I doubt we have more than 10 new doctoral theses a year, which... probably explains why there isn't any great urgency to move to mandated electronic submission - the volumes are a bit low."*

The highest number of PhD theses annually awarded by any respondent from this group was 10. HESA statistics for the academic year 2008-9 record the award of 17,650 doctorates. This suggests that within five years less than 0.05% of theses submitted across the UK will be stored solely in print format.<sup>5</sup>

## 5.2 Digitisation

Respondents with print-only deposit were asked if they digitise theses on demand. 14 do and the rest either do not or left the question unanswered. Of those 14, only 2 digitise in-house and the rest use the British Library EThOS service.

# 6. Open Access

## 6.1 Formal policies and open access

62 of the 70 respondents with formal deposit policies gave details of the default access position in their policy. 90% of current electronic deposit policies are 'opt-out', i.e. theses are made openly accessible by default unless an exception is granted.

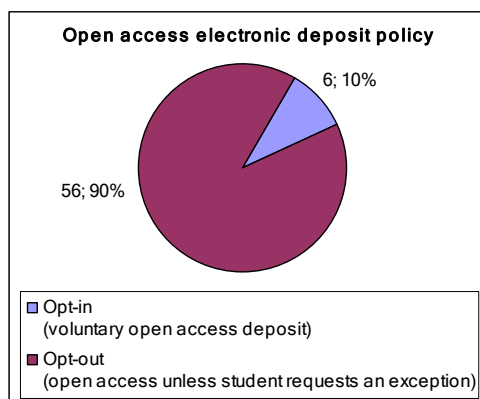


Figure 5: Open access policy

The 6 universities with 'opt-in' policies are from the full range of the HEI spectrum, from small unaffiliated institutions to members of the Russell Group. This suggests that the decision to adopt an opt-in policy is unrelated to the size of the institution.

## 6.2 Making theses available

While the default position at most institutions (81%) is to make thesis metadata available via the library catalogue, policies on access to theses themselves are much more varied.

Closed access print collections of theses are the most likely to be restricted to intra-institutional access. Nevertheless it is clear that open access is overwhelmingly more common than otherwise, regardless of the format or storage medium of the thesis.

<sup>5</sup>[http://www.hesa.ac.uk/index.php?option=com\\_datatables&Itemid=121&task=show\\_category&catdex=3#quals](http://www.hesa.ac.uk/index.php?option=com_datatables&Itemid=121&task=show_category&catdex=3#quals)

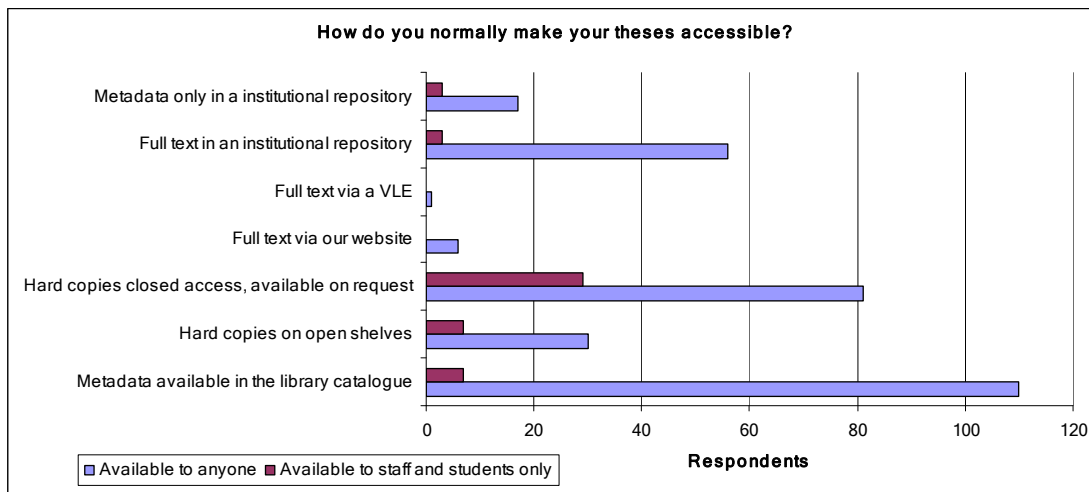


Figure 6: Accessibility of theses

56 institutions responding to this question make full text available in a repository and a further 6 via a website: in total, 62 of the 90 HEIs that currently handle e-theses place them on open access.

This indicates that 69% of current electronic thesis-holding institutions offer open accessibility. As the findings at 5.1 indicate, if all UK HEIs were to implement formal policies along the predominantly 'opt-out' lines of those already adopted in the sector, we might expect electronic theses from at least 90% of institutions to be openly and freely available by default.

Another notable finding is that HEIs with e-theses deposit were four times more likely to store their hard copies off-site, using electronic copies for distribution and sharing.

### 6.3 Repositories and e-theses storage

HEIs with electronic deposit and those who digitise print copies were asked how they store e-theses. Of the 81 respondents, 72 (88%) already have an institutional repository (IR). Of the respondents without repositories, only two indicated that they do not intend to implement an IR in future.

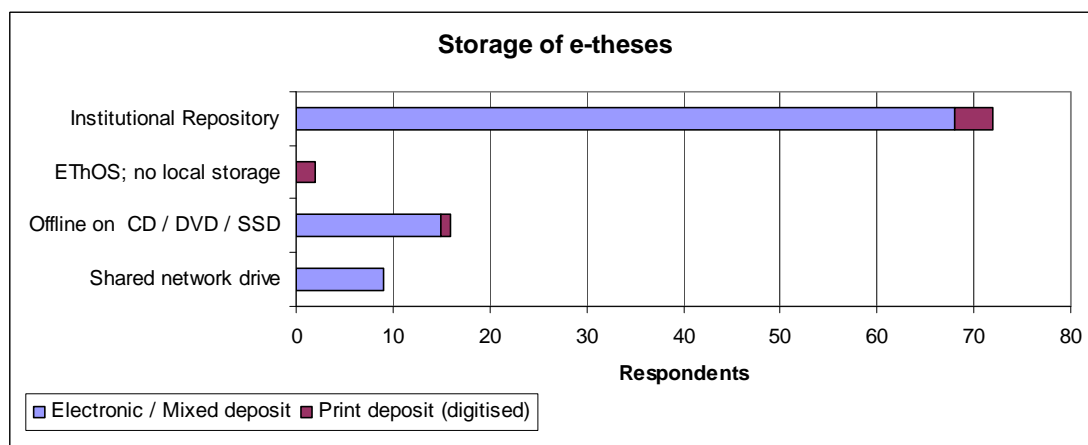


Figure 7: Storage of e-theses

The use of offline or network drive storage may be in addition to inclusion in the IR: several respondents noted that such storage is used for restricted and embargoed theses.

## 6.4 Restricting access to theses

Respondents were asked for the source of access restrictions applied to their theses.

50% of all respondents reported institutionally imposed restrictions. In addition, 90% of those with e-theses allowed authors to impose restrictions on access to their work.

It is noteworthy that restrictions on grounds of 3<sup>rd</sup> party copyright, data protection or potential risks to personal safety were reported only amongst e-theses and not in the print environment.

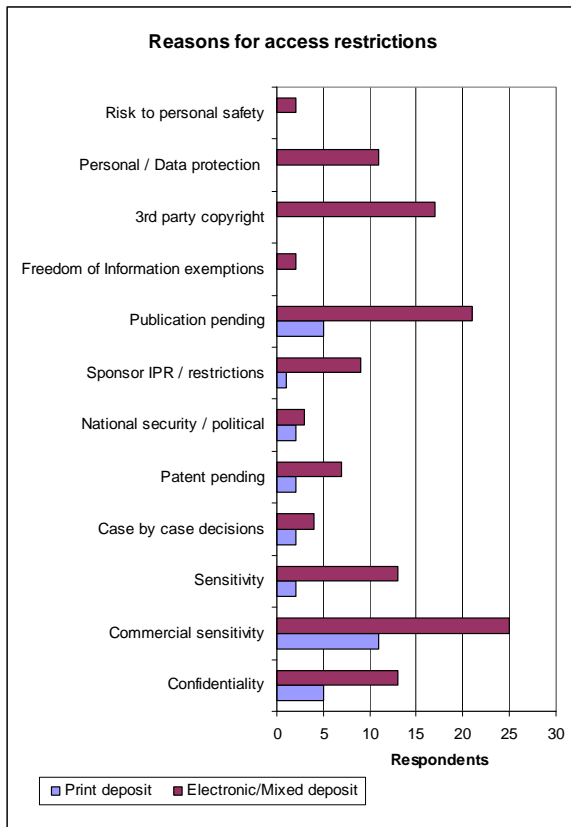


Figure 8: Access restrictions to theses

For both author- and institution-imposed restrictions, the only consistently adopted measure across the sector is the embargo, used by 14 (26% of 54) respondents with print-only deposit and 55 (61% of 90) respondents with electronic/mixed deposit.

The nature of the embargo varies greatly from automatic:

*“Automatic year-long embargo on PhD theses to allow students time to publish”*  
(1994 Group university, 3-5000 doctoral students)

to straightforward:

*“On request a one year, repeatable restriction can be applied. No authorisation from supervisor/head of department required”*  
(Russell Group university, 1-3000 doctoral students)

to exceptional:

*“Our Research Degree Board will embargo the complete thesis if that has been requested by the student and their supervisor. Reasons would have to be very good for this to happen i.e.*

*funded by an external organisation and commercial interests are linked to the work”*  
(University Alliance university, fewer than 1000 doctoral students).

Default periods for embargos vary between 1, 2 or 3 years, and these are frequently extendable at the request of student or supervisor. A more detailed examination of embargo best practice is available on the project website<sup>6</sup>.

## 7. Policies and Practices

### 7.1 Formal policies

Those institutions with electronic or mixed deposit are twice as likely to have formulated policies on areas such as copyright and takedown, and are twice as likely to have made their policies publicly available online.

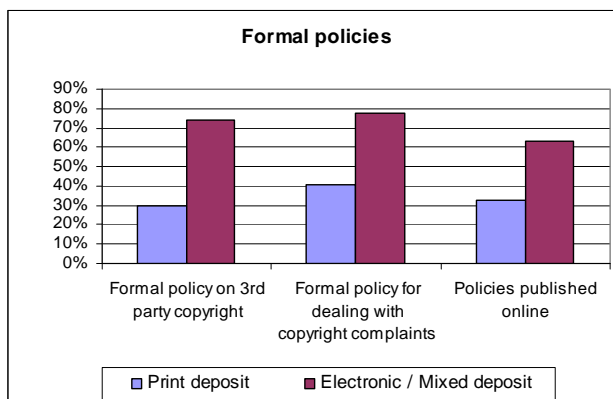


Figure 9: Policies related to theses

### 7.2 Workflow analysis

Respondents were asked to describe their thesis processing workflow. Print procedures were fairly homogeneous. A clear majority followed a version of this procedure:

- 1) Two print copies handed in to a central administrative office.
- 2) One hard copy sent to the academic department.
- 3) One hard copy sent to the library for cataloguing and storage.

Electronic theses workflows, in contrast, are diverse with no clear consensus, possibly due to the complexity of requirements for publication permissions and managing storage and embargoes. For example e-theses may be uploaded to the repository by library staff, faculty administrators, Graduate Office staff or by the authors themselves. The electronic deposit model has flexibility but this poses problems for institutions wishing to streamline or standardise workflows and procedures, as raised in [section 8.6 below](#).

The most common workflow follows this generic pattern:

- 1) Author deposits thesis at a central department.
- 2) Thesis is passed to the library.
- 3) Library staff archive thesis in an institutional repository.

<sup>6</sup> Project website <http://www.ucl.ac.uk/ls/etheses>

Additional stages may be required for internal communications, checking copyright, metadata creation, embargo procedures and the management of hard copies. A more detailed examination of e-theses workflow is available on the project website<sup>7</sup>.

### 7.3 Collaboration

Respondents were asked about present or planned collaboration and co-operation with e-theses. A high level of collaborative activity was indicated by the 122 responses.

- 99 respondents (81%) noted membership of the British Library’s electronic thesis service, EThOS. Of these, 3 also mentioned their own repository’s theses being harvested into EThOS.
- 10 respondents mentioned DART-Europe.
- 8 respondents are involved with the Welsh Repository Network and its collaboration for harvesting e-theses into the National Library of Wales.
- Other local consortia mentioned are SHERPA-LEAP (University of London) and White Rose (the Universities of Leeds, Sheffield and York).

### 7.4 Metrics

Respondents were asked if they keep metrics on the number of theses added, views of theses collection as a whole, views of individual theses or downloads of individual theses. 64% of respondents count the number of theses added; a predictably high figure since the metric is equally applicable to print and electronic theses. At present, more HEIs have plans to record access statistics than mechanisms in place to do so.

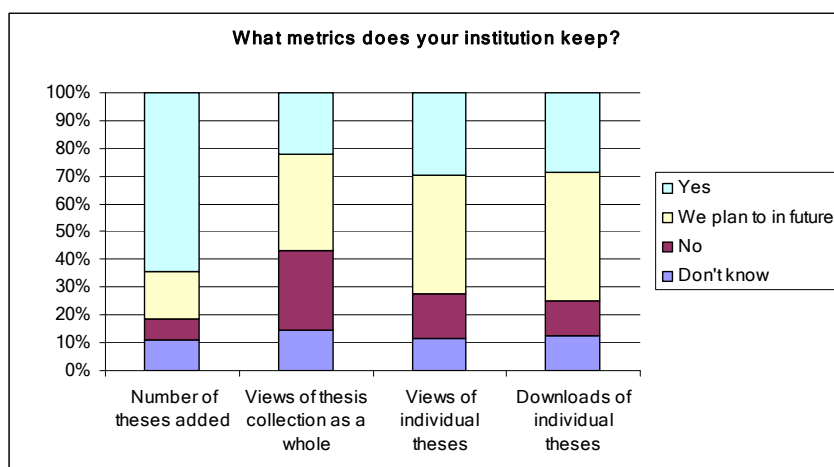


Figure 10: Metrics

Respondents who answered “Yes” to all, and who gave permission, were approached for more detailed metrics. The following response, from a Russell Group university, clearly spells out the high performance of e-theses in terms of access, especially compared with print theses.

- Approx 200 e-theses added per year.
- Approx 200 downloads per year for an “average” thesis
- Compared with print theses: “Anecdotal evidence from the ILL department [which manages all thesis consultation requests at the institution] suggests the most popular theses were requested up to 10 times per year”

<sup>7</sup> Project website <http://www.ucl.ac.uk/ls/etheses>

The download rates for the most popular theses from this institution ranged from 1,494 to 644 for the five months from January to May 2010. The top three theses were each downloaded more than 1,000 times.

Comparison of e-theses with other material in repository (Jan-May 2010)			
Item type	No.	Downloads	Downloads /item
Working Paper	78	10,313	132
Preprint	136	17,944	131
PhD Thesis	1196	121,754	101
Book Chapter	90	9,168	101
Report	46	4,274	92
Journal Article	714	57,901	81
Technical Report	60	3,192	53
Conference Paper	385	17,918	46
Other	91	3,495	38

Another university, unaffiliated to any group, which adds an average 120 doctoral theses per year, reported that only 20% of access to e-theses was from users within their institution, whilst 80% of access originates from elsewhere in the UK and overseas.

This HEI also noted that the “average” thesis has been downloaded 1790 times since Feb 2008, but cautioned that web spiders and crawlers may well account for 90% of the traffic. Taking this into account the average thesis has been downloaded an estimated 179 times since Feb 2008.

## 8. Barriers to e-theses deposit

Research<sup>8</sup> suggests that the primary concerns surrounding electronic, open access theses are:

- Impact on future publication
- Copyright
- Plagiarism
- Sensitive content

Evidence was sought on each of these issues in order to identify the extent and origin of the concerns. Respondents were asked to indicate from which group within their institution each concern tended to emerge: students, supervisors or administrators. Respondents were also asked to describe any actions taken to mitigate each concern. In order to ensure a representative picture of the kinds of mitigating actions taken across the sector, answers here were collected as free text and then classified to identify areas of consensus. Findings are summarised here; a more detailed examination of the responses to each concern is available on the project website<sup>9</sup>.

<sup>8</sup> Literature review, <http://eprints.ucl.ac.uk/20424/>

<sup>9</sup> Project website <http://www.ucl.ac.uk/ls/etheses>

## 8.1 Impact on Future Publication

Respondents were questioned about the perceived concern that making a thesis available online may have an impact on students' opportunities to publish their work. Of 130 HEIs answering the question, 72 (55%) are aware of it being a concern in their institution.

Students are the most likely to express concerns about publication. Anecdotal evidence from the qualitative sample suggests that this is most likely in arts and humanities disciplines. The most substantial concerns presented by respondents were:

- "The publisher on request made it clear that deposit in the IR was not acceptable according to their publishing agreement."
- One respondent quoted a supervisor as having said "I want my students to be able to publish in journals and this is prior publication"
- "[A named University Press] has been approached about open access publishing and expressed the concern in a hypothetical context."
- "All indicated that 'they had heard' that publishers may not want to publish their work if their thesis was available online."
- "We have been asked to remove a thesis which the student hopes to publish as a book."

Only a total of 10 HEIs (7% of respondents) report that this is a frequent or top concern.

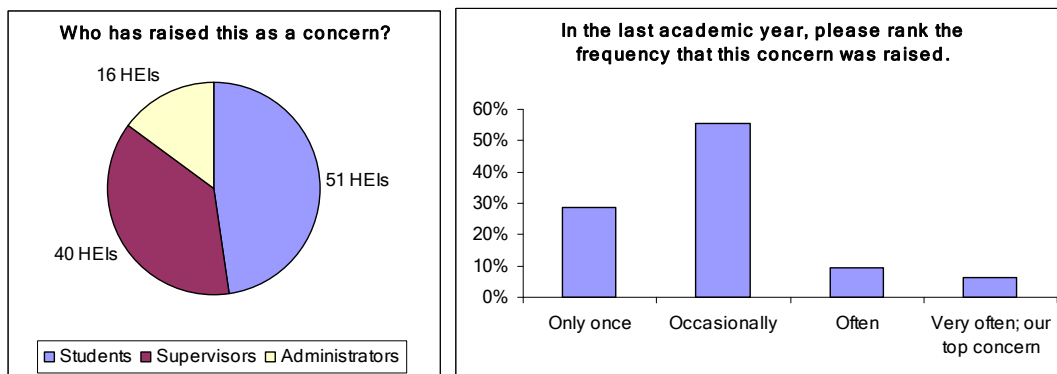


Figure 11: Impact on future publication: Number of HEIs where concern raised

Figure 12: Impact on future publication: Frequency of concern

Of those who took action to mitigate this concern, the majority (59%) named embargoing theses as a solution. Additional access restrictions were mentioned by a further 21%, such as a takedown policy, allowing 'opt-out' from open access, or publishing metadata and abstract only. The second most common measure (50% of respondents) was built around advocacy and educating stakeholders by providing guidance, websites, or other information. Some universities recommend that students check with their publisher or use OA-friendly publishers; others suggest that OA theses may even attract publishers.

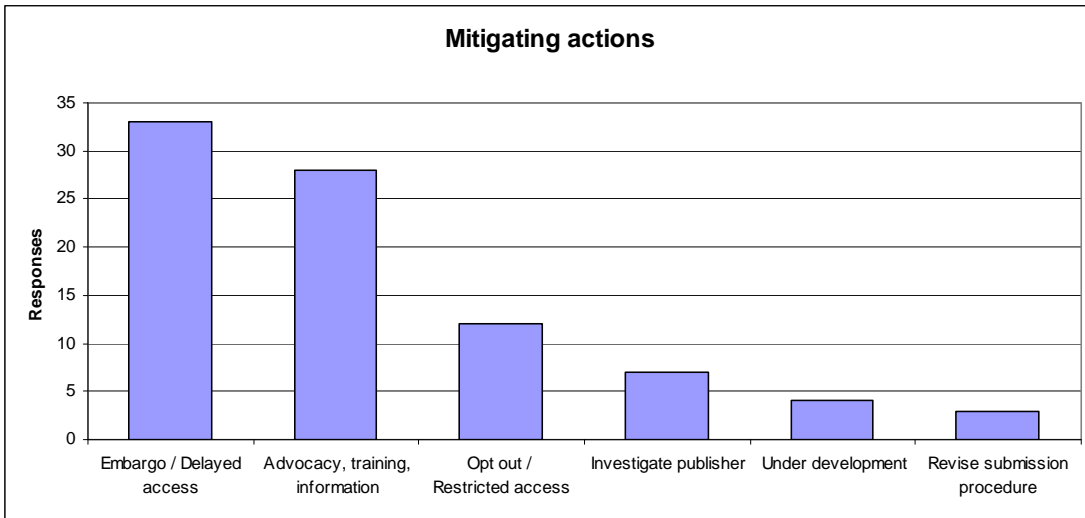


Figure 13: Impact on future publication: Mitigating actions

Notable comments:

- “Most students and supervisors don't understand these areas and are confused/worried. Normally when there has been an opportunity to explain the worries fade”
- “Publishers who are genuinely interested will still consider PhDs to publish.”
- “Open access policy to thesis publication is a given; students may be reassured by the awareness that the data and text of a thesis may need to undergo considerable review prior to publication”
- “We spelt out the positive benefits of making the theses available and in fact some research suggests that publishers are more likely to publish theses if they are being well accessed online”
- “We also have students who published theses and were given permission by their publishers to deposit e-theses”

When asked to quantify incidences of publication refusal, no concrete examples were offered.

- A Russell Group university commented: “None (or for the previous 6 years we have been putting e-theses online)”

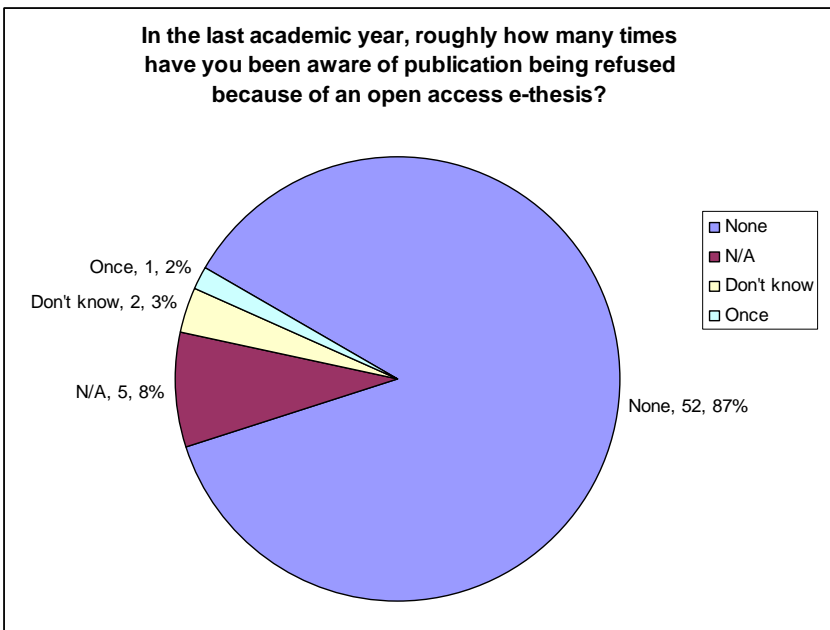


Figure 14: impact on future publication: Actual occurrences

Five respondents pointed out that an effective embargo system prevents the issue from arising.

The single instance reported could not be substantiated:

- “I heard of one case indirectly re. a monograph publication but not the precise details and don't really know the accuracy of the story”

## 8.2 Third Party Copyright

89 respondents (62%) are aware of concerns within their institution surrounding IPR and third party copyright in theses.

The issues cross many academic disciplines. Examples of concerns raised:

- “Owners of 3rd party copyright may sue us”
- “Theses in arts subjects including lengthy quotations from authors under study; theses in science subjects including diagrams, plans and maps.”
- “All art history theses are awash with images, most of which need copyright clearance”
- “Copyright law does not define 'a short quotation' so students occasionally have concerns about exactly what they must seek permission for.”
- “Difficulties in covering this for retrospective theses where we can no longer get a declaration form from the author to cover copyright concerns.”
- “The issue of having to contact people to obtain copyright is seen as onerous and unfair to expect this work to be performed by the student.”
- “Not sure that students and supervisors are aware of the implications of including third party copyright material in what is essentially a published work.”

Adherence to copyright is primarily a concern for administrators, who are responsible for making the work available. A total of 22 respondents (15%) report that this is a frequent or top concern.

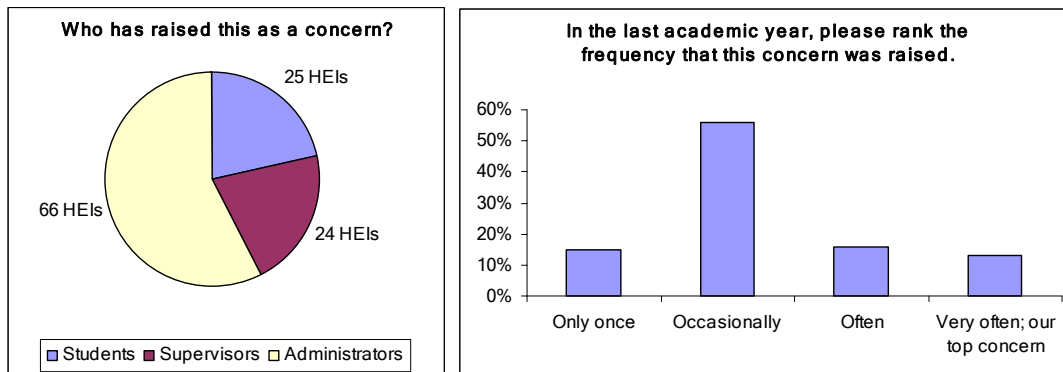


Figure 15: Third party copyright: Number of HEIs where concern raised

Figure 16: Third party copyright: Frequency of concern

For those taking action to address the concern, 70% of measures revolve around teaching good practice and provision of guidance, including templates for permissions requests. 16 respondents (21%) mention their takedown policy. 13 HEIs restrict access for example via embargo or opting out altogether. 12 HEIs mention checking the thesis, variously by the student, repository staff, subject librarians or the Rights department. Only 11 HEIs advocate creation of a separate digital edition with copyright material removed.

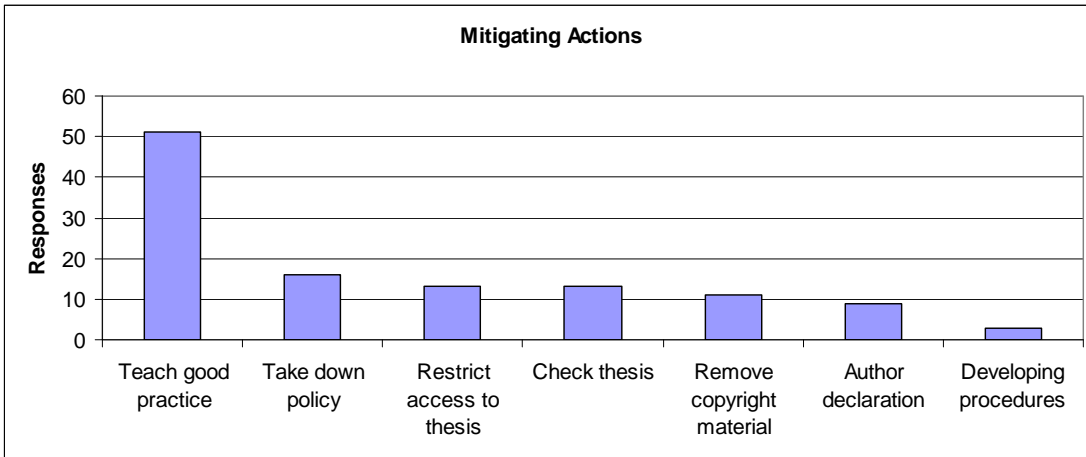


Figure 17: Third party copyright: Mitigating actions

Notable comments:

- “We strongly suggest that a good copyright habit is a good academic skill and goes beyond theses.”
- “Exhaustive education and training, thorough checking, tight policies and a rigorous Take Down policy”
- “Seek publisher permission as you go along - don't leave until end”
- “Support reasonable and informed risk.”
- “This is a relatively low risk educational environment but not to be complacent.”

When asked to quantify incidents where action had to be taken to remove access to a thesis, 64% of respondents had none to report, with two respondents attributing this to the success of their system/procedures. Eight HEIs had between 1 and 3 theses removed. Only three HEIs reported higher figures (12, 15 and 50).

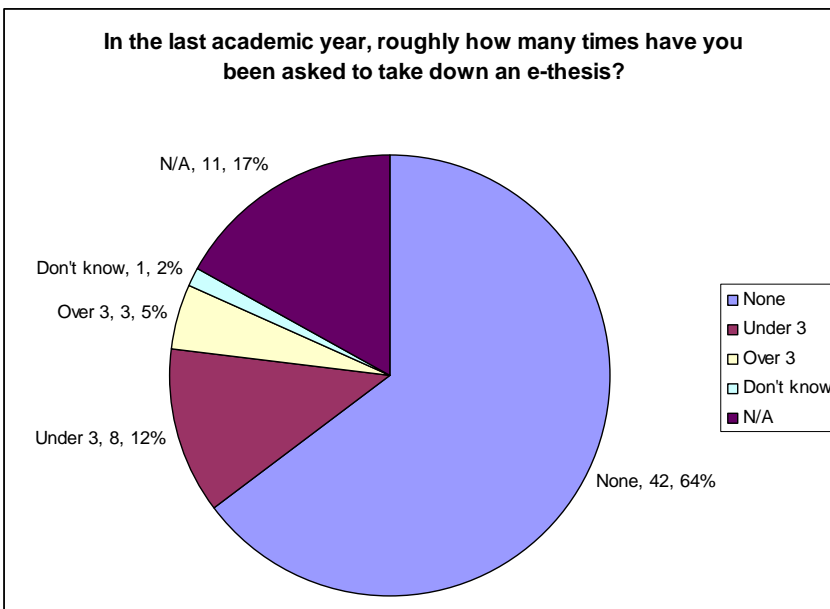


Figure 18: Third party copyright: Actual occurrences

### 8.3 Plagiarism

76 respondents (53% of HEIs) reported concerns about plagiarism. The concerns raised are typified as “General fears that open access means free for all on student's IPR” or specific queries such as

“Can we ‘lock’ thesis files to prevent cut and paste”. One respondent believed that concerns are “decreasing as students gain confidence in the value of the open access IR”.

Plagiarism is a concern for almost equal proportions of students and supervisors. Only 5 respondents (3% of HEIs) report this as frequent and none say plagiarism is a top concern.

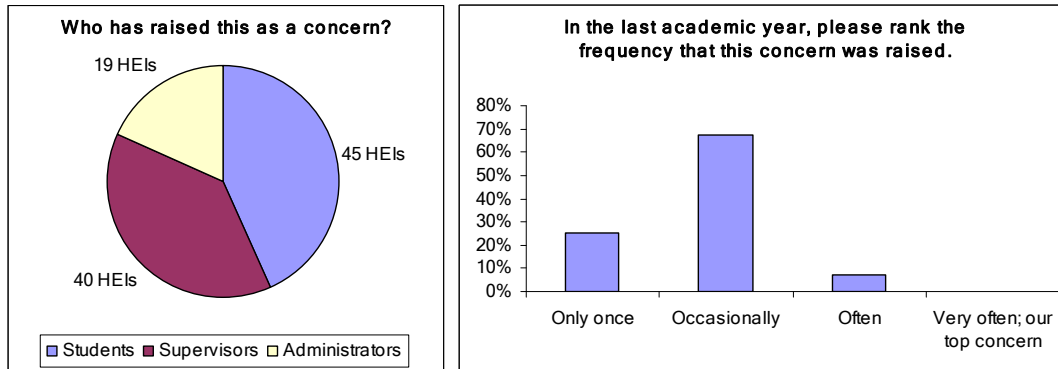


Figure 19: Plagiarism: Number of HEIs where concern raised  
 Figure 20: Plagiarism: Frequency of concern

Of those taking action to address these concerns, a total of 64% (38 HEIs) concentrate on advocacy and education, including 23 who suggest that OA actually assists plagiarism detection. Four pointed out that there is a risk of plagiarism with print theses too. Almost 20% of those answering the question (12 HEIs) also advocate detection software as a means of identifying plagiarism in authors' work and thus protecting the institution. Some universities prefer to restrict access via embargo or opting out. Creative Commons licensing is mentioned, and one points out that users of EThOS need to register and identify themselves to allow access.

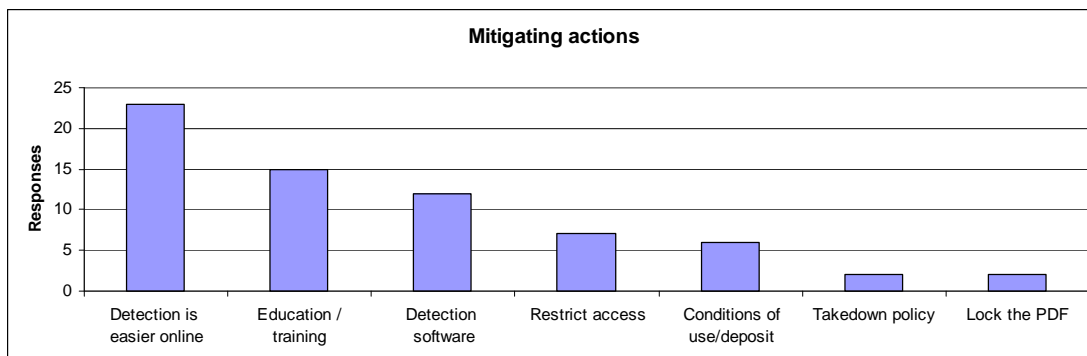


Figure 21: Plagiarism: Mitigating actions

**Notable comments:**

- “Deposition serves to “date-stamp” the author's claim to priority”
- “Easier to detect if work has been copied from an online source than from a hard copy hidden in a basement”
- “We advertise such services as TurnItIn and do not shy away from fact that copying is a real risk.”
- [print copy] “Restrict any copying, but not access or brief quotations”
- “It was felt by the relevant committee that the benefits far outweighed the concerns”

When asked to quantify instances of plagiarism, very few were reported. Of those answering the question, 44 (67%) knew of none, five knew of one incident and only one reported two.

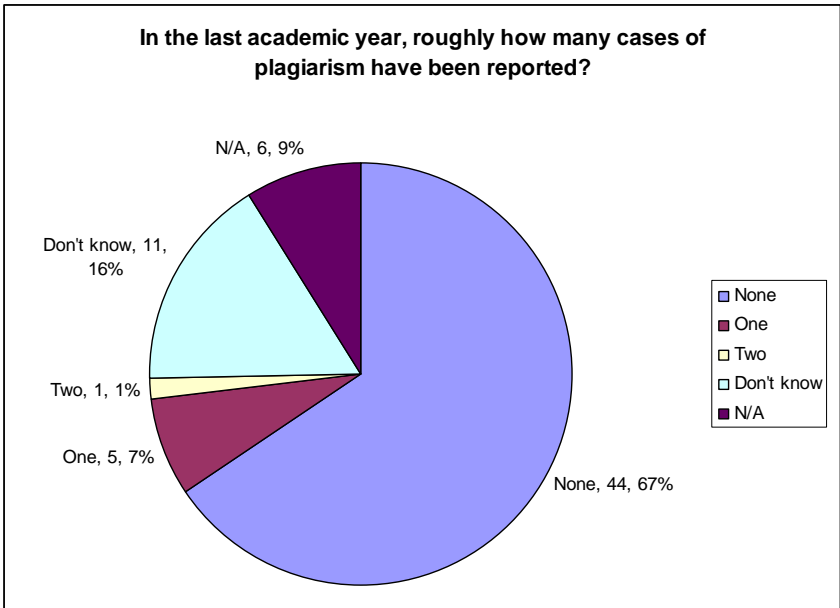


Figure 22: Plagiarism: Actual occurrences

### 8.4 Sensitive Content

75 respondents (52% of HEIs) report concerns about sensitive data within theses. The nature of sensitivity is varied:

- “Commercially sensitive information”
- “Theses have contained data that is a) relevant to court case, b) contains animal experimentation, or c) contains information on child welfare cases that could be identified in the field, even though anonymised.”
- “Intelligence and security theses”
- “Students inform us that participants in (eg. case study) theses often only agree to be interviewed if confidentiality is guaranteed”
- “Health theses contain patient information”
- “Author considered life would be at risk”

Concerns about sensitive data are equally divided between students, supervisors and administrators. Only 3 respondents (2% of HEIs) reported this as a frequent or top concern.

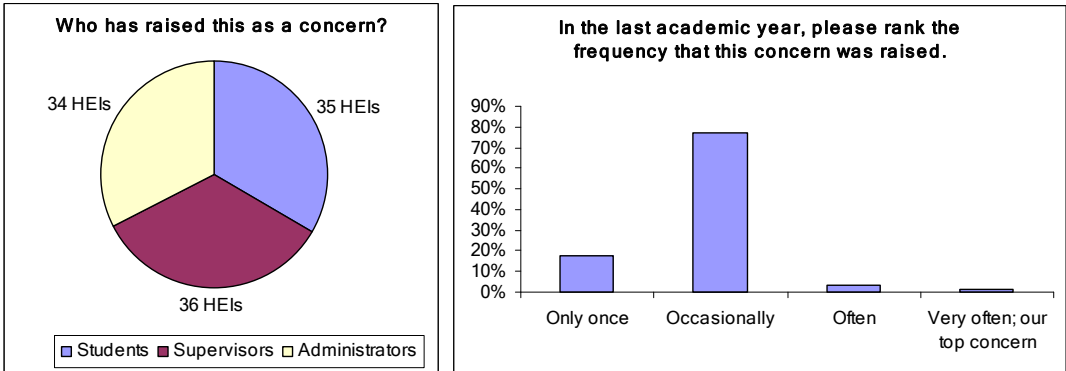


Figure 23: Sensitive content: Number of HEIs where concern raised

Figure 24: Sensitive Content: Frequency of concern

Of the respondents taking action to address the concern, the vast majority (79%) restrict access to the thesis, 47% via embargo and 32% via other restrictions e.g. opt-out altogether. Only 17% advocate

the creation of a different or redacted version for public access. Instead, 25% use education and training or expect candidates to omit sensitive data from their theses.

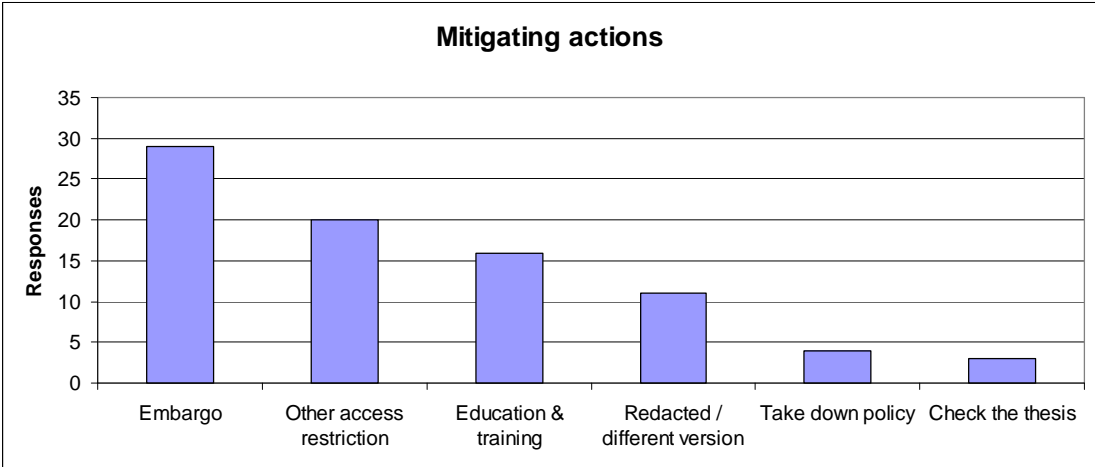


Figure 25: Sensitive content: Mitigating actions

Notable comments:

- “Removing sensitive material to an appendix and embargoing that separately so that the main body of research is still available publicly”
- “Students should be following good practice in these areas in any case (regardless of electronic access).”
- “Students should notify their subjects that the output from their study may be placed in a digital archive, which is available globally.”
- “Research inappropriate if none of it can be made publicly available or not available at all”
- “Standard ethical declaration form and referral to university ethics committee when required.”

When asked how many theses require embargo or redaction, only 29% of HEIs said “none”, which is lower than with other concerns. Another 29% had fewer than 5 affected theses, and a further 10% up to 40. Only 2 HEIs report a higher number (300 and “a significant number”, both high volume Russell Group universities). Of those who said “don’t know”, three noted that the process is pre-library and they would therefore not receive or know about those theses.

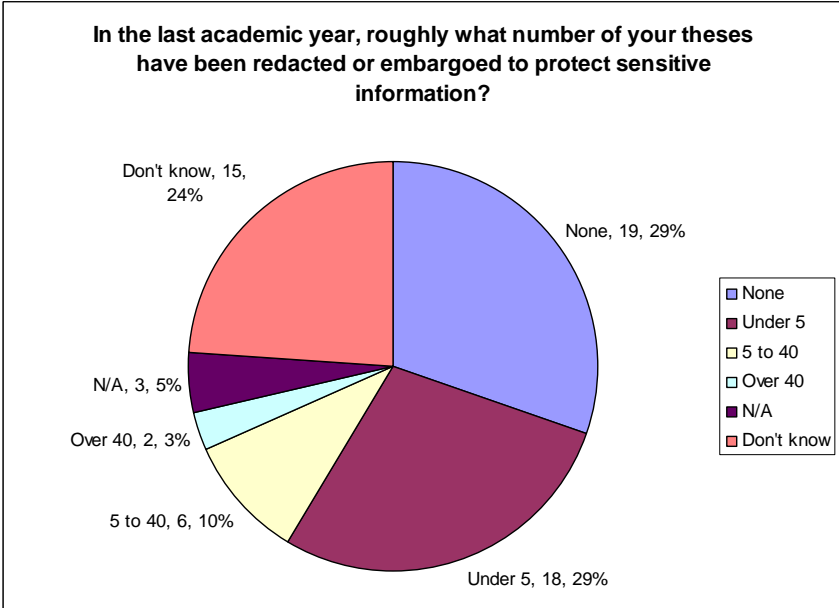


Figure 26: Sensitive content: Actual occurrences

## 8.5 Practical barriers

In addition to earlier concerns discussed, respondents were asked whether they have encountered practical barriers to deposit and open access of e-theses.

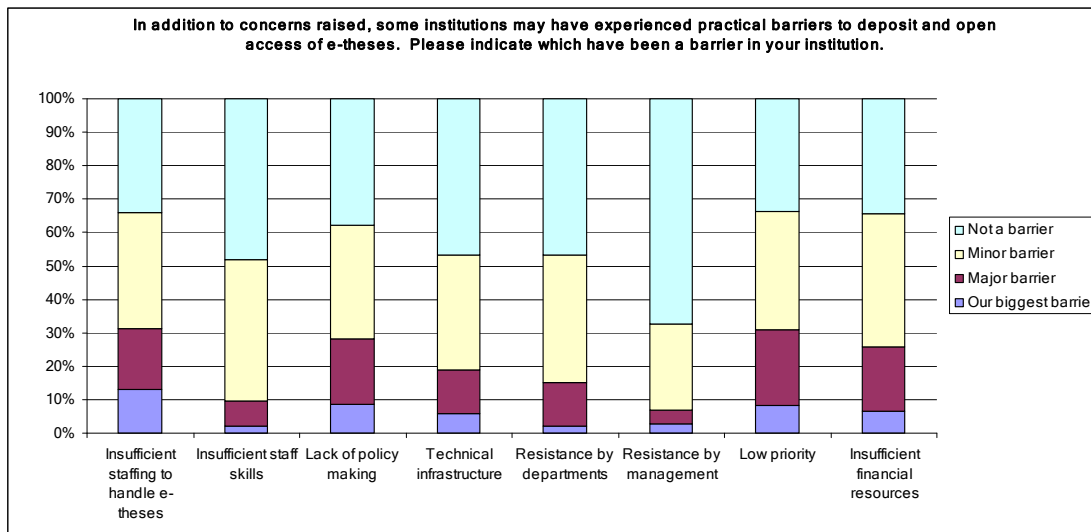


Figure 27:

Practical barriers to e-thesis deposit: percentage comparison

None of the potential barriers was cited as a major or top barrier for more than 31% of respondents. Resistance by management and departments is very low: in particular resistance by management was reported as “not a barrier” by 70 respondents. However, lack of priority and policy making are comparatively high, with lack of priority cited as a major or top barrier by 31 HEIs. This suggests that while acceptance and willingness for OA e-thesis deposit may be prevalent, some institutions still lack the impetus to implement it.

Not surprisingly the highest figure for “our biggest barrier” is a resourcing issue, that of insufficient staffing to handle e-theses. Financial resources and technical infrastructure are seen largely as only a minor barrier or not an impediment at all, and staffing skills are not seen as a barrier.

## 8.6 Other concerns

Respondents were invited to explain any other concerns they had encountered and what steps were taken.

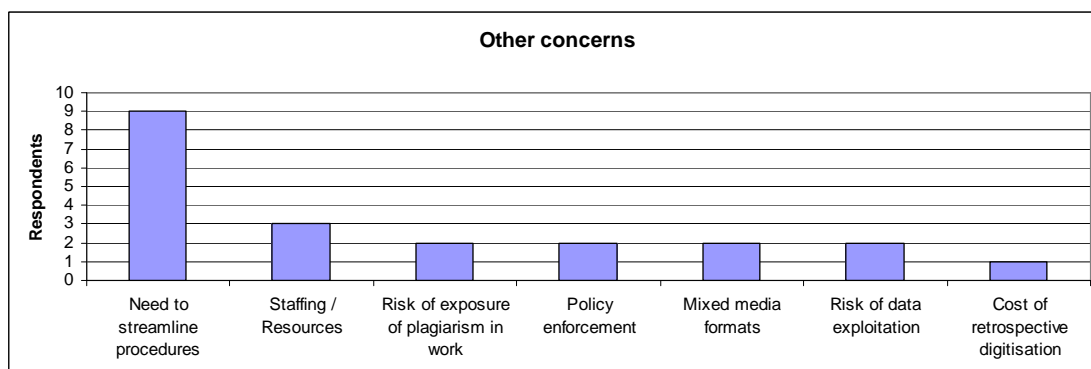


Figure 28: Other concerns on e-thesis deposit

The need to streamline procedures is the largest concern, raised by nearly 40% of those answering. This is borne out by the diversity of processes in raw data responses to the Workflow question in this survey ([section 7.2 above](#)). Five respondents say the issue is ongoing; others report that communication and in one case introducing an IR have been the solution. A further 9% add the need for policy enforcement.

Sample comments:

- “Need for streamlined deposit process joining up separate IT systems”
- “Relationship of supervisor to thesis. Although the student is the copyright owner, supervisors have vested interests in publication from theses too, and often in negotiating commercial sponsorship”
- “Raised mandated e-submission back in 2001 and it has taken until 2009 to get it back on the table. Slow process.”
- “Examined and electronic thesis must be identical... We've had to reflect it in our policy”
- “Actually e-theses are in the bigger picture far less important than published output by staff”

Some concerns echo issues already raised, eg staffing/resources and risk of data exploitation, for which embargo is cited as the solution. Two mentioned concern about risking exposure of plagiarism in students' work; both cite Turnitin as the solution.

Non-text media are an emerging concern for two HEIs offering doctorates in the musical and creative arts. One notes:

*“Practical concerns over who, how and to what standard are non-text PhD work recorded and converted to readily accessible on line formats... What happens for say a fashion show or art exhibition? Do you make the student record the event/work in a certain format (still images/video etc). What about the quality? ...Resourcing implications if staff and equipment has to be available for student project/PhD work”*

Not all comments expressed problems. A 1994 Group university about to adopt electronic mandate in 2011 reported “Very few concerns so far apart from two people requesting embargos”.

## 9. Conclusion

The aim of this project was to gather evidence to encourage HEIs to prioritise changes to their policies and practices that result in electronic versions of their new PhD theses being available on an open access basis from their institutional repositories. The survey results show that present and planned electronic deposit in UK HEIs is already very high, and that the 13% of UK HEIs still planning to remain with print deposit represent a very small number of PhD theses produced in the UK.

Perceived barriers to electronic deposit were examined, and shown either to have minimal negative outcome or to be successfully managed with proven strategies across the sector.

The accessibility of the current e-thesis corpus, of which an estimated 69% is open, could be improved. Most of the formal policies that were reported involve ‘opt-out’ open access; the encouragement of similar policies would only increase accessibility. Sharing good practice in overcoming the perceived ‘barrier’ issues, and making tried and tested workflows, mitigation strategies and model policies more readily available between HEIs, may help to build confidence in all aspects of e-thesis management across the sector, ease the way to formal, open access-friendly policy adoption, and ultimately increase the accessibility of the UK's electronic research theses.

## 10. Recommendations

### *Measures to increase open access to UK e-theses*

1. HEIs should be encouraged to adopt formal, OA-friendly 'opt-out' policies. To help to achieve this, work should be carried out to help HEIs to share experience and build confidence in dealing effectively with the most common concerns around e-thesis management, facilitating the sharing of good practice, successful workflows, and mature and proven policies.

2. The project found no hard evidence that open access thesis deposit prejudices an author's prospects of publication. Follow-up work to investigate publisher attitudes could be undertaken; the creation of a central knowledge base (or the extension of SHERPA/RoMEO) might be appropriate.

### *Impact*

3. The ability to demonstrate the impact of open access theses is potentially beneficial to institutions. Few HEIs maintain relevant statistics; there is inconsistency of method among those that do. Standard metrics for e-thesis access should be agreed; repository plug-ins for the most common platforms developed; and guidance on the effective use of third-party resources such as Google Analytics and Google Webmaster Tools should be collated and shared.

### *Looking ahead*

4. HEIs are beginning to move away from archival thesis storage in print form: 8 HEIs already no longer accept theses in hard copy. It would be timely to explore strategies for helping HEIs to meet the challenges of the long-term digital curation of electronic thesis collections.

5. The quantity of non-text theses, such as multimedia submissions for performing arts degrees, and supplements to text such as datasets, simulations, etc., is likely to increase as e-deposit rates grow and the availability of enabling technology continues to expand. Investment in co-ordinated 'best practice' resources to support the sector in meeting the challenges of managing non-text theses – issues such as format selection, metadata standards, preservation strategies, and so on – should be considered.