

Plagiarism

In June 2010 UCL Library Services conducted a sector-wide survey into thesis deposit and open access in UK universities and HEIs. A rich body of data was provided by respondents describing how they approach the issues surrounding open access deposit of electronic theses. The data was analysed and evidence is presented here of consensus and best practice. A survey summary report and full text responses are available on the project website¹.

Of 144 survey respondents, 76 (53%) reported concerns about e-theses being susceptible to plagiarism. Plagiarism is a concern for almost equal proportions of students and supervisors:

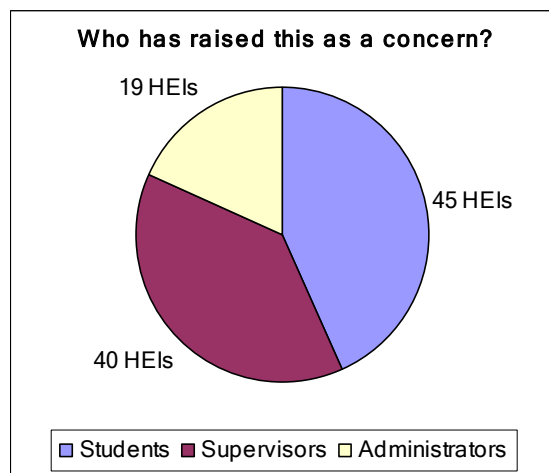


Figure 1: Plagiarism: Number of HEIs where concern raised

Only 5 respondents report this as frequent and none say plagiarism is a top concern.

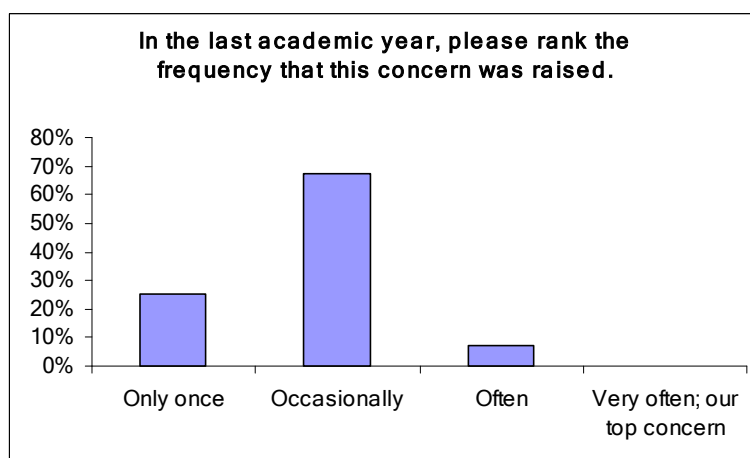


Figure 2: Plagiarism: Frequency of concern

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

Examples of concerns raised

- “General fears that open access means free for all on student's IPR”
- “Can we ‘lock’ thesis files to prevent cut and paste”.
- “Some students are sponsored by companies and worried about patents for their own ideas etc. Reassure them that the existing embargo options remain the same for e version as the print version”
- “One author complained because it was possible to cut and paste from their thesis PDF, which they didn't think should be possible.”
- “Researcher had already been plagiarised in her home country and wanted her thesis removed.”
- “That since the work was openly accessible in an electronic format that it would be easier for others to copy.”
- “They feel threatened in their ownership because of worldwide web access.”
- “Anxiety about re-use of unpublished thesis material without attribution”
- “An isolated case of academic stalking”
- One respondent believed that concerns are “decreasing as students gain confidence in the value of the open access IR”.

Mitigating actions

Respondents were invited to describe in their own words what measures they take to address these concerns. Only three take no action. Of those taking action, a total of 64% (38 HEIs) concentrate on advocacy and education. Notable responses are discussed below.

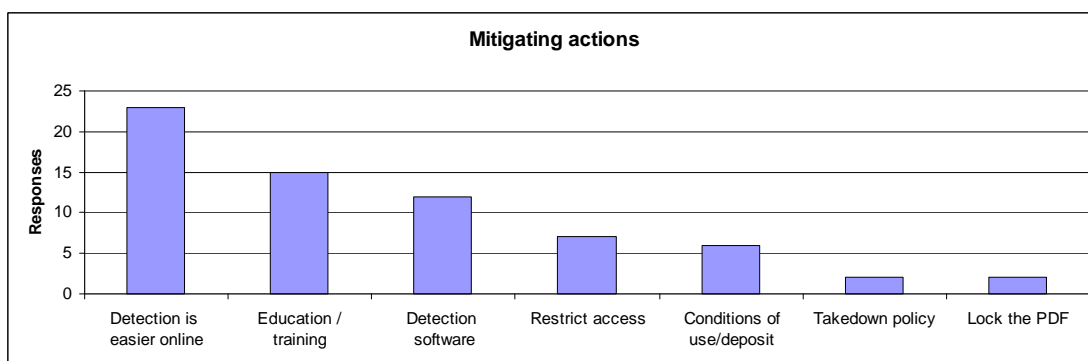


Figure 3: Plagiarism: Mitigating actions

Detection is easier online:

The most common measure is one of advocacy: 23 responses emphasise the fact that OA actually assists plagiarism detection.

- “Gentle pointing out that it’s easier to detect if work has been copied from an online source than from a hard copy hidden in a basement”
- “Our line is that making theses available electronically is actually likely to reduce instances of plagiarism, since any given thesis will be more easy to track down, more exposed to anti-plagiarism tools etc.”
- “Illustrate how much easier it is to detect plagiarism in electronic versions of theses”
- “Deposition serves to “date-stamp” the author’s claim to priority.”
- “One stresses the ease with which any plagiarism can be detected and proven with a date-stamped online publication”

This respondent reports an incident of plagiarism that would not have been detected if the work had not been online:

- “Former student said she found her work being copied in a document by another student.”

Education / training:

The second most popular measure (15 responses) is to provide a range of education and guidance:

- “We do training on the authors’ copyright rights and how the IR emphasises their copyright ownership. “
- “Explain policy and refer to Research Office in needed”
- “Copyright policy is displayed in the Library and on our web pages.”

- “Face to face training, explaining how Google and other search engines, including anti-plagiarism sites, will pick up the thesis”
- “Refer to EThOS guidance”

Several respondents pointed out that there is also a risk of plagiarism with print theses or other published material:

- “Pointing out that the facility to copy already exists in print format”
- We explain how this is also taking place in a print context ([we have] concrete examples) and is difficult to avoid entirely
- The problem is no different from any other e-journal articles or other e-publications the author may produce.

Detection software:

Almost 20% of those answering the question (12 HEIs) advocate detection software as a means of identifying plagiarism in authors’ work and thus protecting the institution.

- “All our theses are checked against Turnitin before being added to the institutional repository. This policy is advertised to students. It’s assumed that the availability of plagiarism software will deter others from plagiarising our students’ work.”
- “Mention that EThOS is working with JISC-iPAS/iParadigms to include theses in the corpus against which suspected plagiarism can be checked.”
- “We advertise such services as TurnItIn and do not shy away from fact that copying is a real risk.”
- “Registry run all theses through Turnitin plagiarism software before sending the e-copy onto the library for deposit.”

This highlights the fact that plagiarism cuts both ways, and there is a real risk to the credibility of HEIs awarding PhDs if they place theses online which contain plagiarised material. Theses submitted electronically are very much easier to check in this way than print copies.

Restrict access:

Seven universities restrict access, for example via embargo or opting out.

- “If the student raises a concern, the abstract only is shown”
- “Reassure them that the existing embargo options remain the same for e version as the print version”

- “We monitor who accesses each thesis (only [our] staff and students may currently have access)”
- “If it remains an issue, remove the thesis.”

However, by blocking access to theses as a result of fears about the misuse of their content, HEIs block legitimate, scholarly use of their content as well.

Conditions of use /deposit:

Six respondents note that their conditions of use for end-users, and deposit declarations by authors, set out responsibilities and restrictions of usage:

- “Clearly state the end-user license terms and conditions of use.”
- “Explained that registration is necessary to use EThOS.”
- “Our repository does have a set of policies which cover permissible re-use and acknowledgement.”
- “The nature of public access to e-theses is explained in a statement which students are required to agree with (by tick-box) when depositing.”
- “We add a cover sheet to e-theses (begun in Feb 2010 with a rolling program to cover older e-theses in progress) setting out fair use and theses author IP rights.”
- “A default Creative commons license is applied which the author may change - Creative Commons Attribution-Non-Commercial-No Derivative Works 2.0 UK: England & Wales.”

Takedown policy:

Some universities employ their take-down policy to address this concern:

- “Policy that authors can request their PhD be taken down if they have sufficient cause to believe their work is being copied /used by another party”

Lock the PDF:

Technical measures such as locking the PDF may deter plagiarists from copying the text of a thesis:

- “Reassurance that there are security settings on each PDF requiring special software to try and unlock it.”
- “If we are sent Word files to convert to PDF, we 'lock' the PDFs to make them harder to edit, copy and paste etc (although if the author sends a PDF, we assume that they are happy for it to be unlocked if that is what they sent).”

If possible though, care should be taken to avoid locking files if that will affect their preservation, by preventing their being migrated to newer formats, for example.

Other measures:

Other responses emphasise the benefits of open access:

- “Explaining reduces concerns. Most students are eager to participate as they see the advantage to researchers.”
- “Often ask the question “Have you used online theses yourself? Benefits to you? Why don't you want other people to see your work and make it easily available to them? Quite often this makes the point.”
- “It was felt by the relevant committee that the benefits far outweighed the concerns”

Actual occurrences

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.

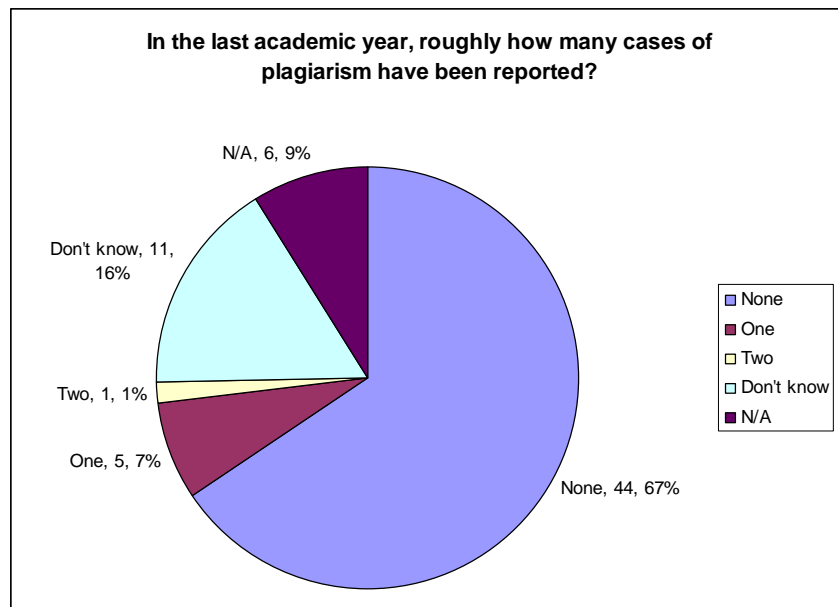


Figure 4: Plagiarism: Actual occurrences

It is worth noting though that plagiarism is more likely to be detected at the plagiarists home institution, rather than by the victim. This emphasises the importance of thorough plagiarism checks prior to award.

Conclusion

While a student or their supervisor might be concerned that their work or research results might be used unscrupulously, and wish to protect their IPR, shutting the thesis away may be a counter-productive measure. The overwhelming majority of users of theses are other researchers engaged in legitimate scholarly endeavours. In this context, the availability of theses is vital for literature reviewers to avoid duplication of efforts and wasted resources.

Authors can be educated to see the benefits of open access, as making available exposes it to plagiarism detection tools and the date of deposit can prove priority. Clear conditions of usage posted before accessing the thesis or the repository can demonstrate that end-users have a responsibility not to misuse the thesis and to acknowledge quotations correctly. Few actual occurrences of plagiarised work have been reported. With a combination of advocacy and clearly stated access regulations therefore, the need for embargo should be minimal.

Authors:

Josh Brown (SHERPA-LEAP) and Kathy Sadler (UCL Library Services), Project Officers on 'Influencing the Deposit of Electronic Theses in UK HE' project.

Further information can be found on the project website at <http://www.ucl.ac.uk/ls/etheses>