



Detecting unauthorised outsourcing of coursework: Metadata in files

Oliver Trevisiol, Christoph Schmitt

Version: 30.05.2023

Outsourcing to third parties, such as ghostwriters or undisclosed text production by artificial intelligence, are considered difficult to detect. Metadata in electronic documents can provide clues at various points. A single indication of potential outsourcing is not yet proof, but several can provide an illuminating overall picture. Further evidence can be provided, for example, by conspicuous features in the bibliography or in the handling of literature.

Note for further use:



This work is licensed under a Creative Commons Attribution 4.0 International Licence.

You are free to share, copy and redistribute the material in any medium or format. You are free to adapt, remix, transform, and build upon the material for any purpose. You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

Additional information about CC licensing:

<https://creativecommons.org/licenses/by/4.0>

Detecting unauthorised outsourcing of coursework: Metadata in files

Metadata on DOCX, ODT or PDF files may contain evidence of contract cheating or other unauthorised help by third parties or artificial intelligence. Individual items of data are not certain proof, but the overall picture can provide good evidence, as unauthorised help is often sought in a time crunch and no effort is then made to manipulate metadata. However, it is of course possible to delete such data. It is also possible to create a submission version in a new file. Both are perfectly legitimate.

I. MS Office

Properties / Total editing time

A very low value suggests that little work has been done with the document.

Properties / Company

Does the entry contain an explicable name such as that of one's own university? The field may contain references to a contract cheating company or an essay mill.

Related Dates / Created

Date of creation of the file. Is the date plausible?

Related Dates / Last modified

Date of the last modification of the file. A very short interval between the creation date and the last modification date means that the file was only worked on for a short time.

Related Dates / Last printed

Date of the last printout or creation date of a PDF file.

Related People / Author

Contains the user name entered in the Office Options. Does the entry contain the student's name or another explainable name? May contain references to a contract cheating provider.

Related People / Last modified by

Contains the user name entered in the Office Options. Does the entry contain the student's name or another explainable name? May contain references to a contract cheating provider.

File / Properties (Right-click on file in the file manager) / Details / Origin / Revisions number

The revision number counts the number of manual saving operations. A very low number suggests that little work has been done with the document.

File / Properties (Right-click on file in the file manager) / Details / Origin / Total editing time

A very low number suggests that little work has been done with the document.

II. LibreOffice

File / Properties (Right-click on file in the file manager) / Created:

Date of creation of the file. Is the date plausible?

File / Properties (Right-click on file in the file manager) / Modified:

Date of the last modification of the file. A very short interval between the creation date and the last modification date means that the file was only worked with briefly and possibly with copy and paste.

File / Properties (Right-click on file in the file manager) / Statistics / Author

This field contains the user data stored in the LibreOffice options, in this case name and / or company.

File / Properties (Right-click on file in the file manager) / Statistics / Modified

Date of the last modification of the file. A very short interval between the creation date and the last modification date means that the file was only worked with briefly and possibly with copy and paste.

File / Properties (Right-click on file in the file manager) / Statistics / Revision number

Counts the number of manual saving operations. A very low number indicates that little work has been done with the document.

File / Properties (Right-click on file in the file manager) / Statistics / Total editing time

A very low number indicates that little work has been done with the document.

III. PDF files

PDF files contain less metadata than Office files. The data may also vary depending on the software used to create it. The use of professional software that is paid for and not licensed to the university can be an indication of a contract cheating provider.

File / Properties / Author

Takes the user name from the Office programme. Can also be entered separately.

File / Properties (Right-click on file in the file manager) / PDF Information / Author

Takes the user name from the Office programme. Can also be entered separately.

References

The composition is based on:

Crockett, Robin (2022): Presentation, Properties and Provenance. The Three Ps of Identifying Evidence of Contract Cheating in Student Assignments. In: Sarah Elaine Eaton, Guy J. Curtis, Brenda M. Stoesz, Joseph Clare, Kiata Rundle und Josh Seeland (Hg.): Contract Cheating in Higher Education. Global Perspectives on Theory, Practice, and Policy. Cham: Palgrave Macmillan, S. 169–184.