# Removal of sensitive data from documents

03.07.2024 13:25:57

**FAQ-Artikel-Ausdruck**

| | | | |
|---|---|---|---|
| **Kategorie:** | Datenschutz | **Bewertungen:** | 0 |
| **Status:** | öffentlich (Alle) | **Ergebnis:** | 0.00 % |
| **Sprache:** | en | **Letzte Aktualisierung:** | 10:31:36 - 27.01.2023 |

***Lösung (öffentlich)***

When creating Office and PDF documents, in addition to the directly visible content, various so-called metadata is also stored, including the names of authors, file paths or creation time stamps. This also applies to all further embedded documents, e.g. inserted images, which in turn have their own capacity for storing metadata. In order to prevent the disclosure of sensitive and personal data, the recommended procedure differs depending on the type of document:

Microsoft Office

For documents created with Microsoft Office 2016 and newer (Word, Excel, PowerPoint, etc.), the Central Data Protection Office of the Universities in Baden-Württemberg (ZENDAS) maintains a website [1]Hidden Data in Microsoft Office 2016 Documents (accessible outside of the campus network only per [2]Shibboleth login via DFN-AAI) as an overview of potentially sensitive data contained in the documents and options for deletion. Specifically in Microsoft Word, Excel, PowerPoint and Visio, most hidden data can be removed using the built-in Document Inspector. ZENDAS  provides [3]detailed instructions for this. Microsoft also maintains its own instructions for using the Dcument Inspector in the support article [4]Remove hidden data and personal information by inspecting documents, presentations, or workbooks.

PDFs

PDF files can contain metadata and embedded documents that need to be checked for sensitive data before sharing. Adobe provides detailed instructions for [5]Removing sensitive content from PDFs. However, a license of Adobe Acrobat Pro is required to execute the steps mentioned there.

As a free and privacy-friendly alternative, you can use the [6]PDF24 Creator, which provides rudimentary tools for redacting and removing metadata from documents. Please use this only via the version that can be downloaded directly to your computer and refrain from editing sensitive documents in the cloud.

To assist with the automated cleanup of potentially sensitive metadata, TUD-CERT provides the webservice[7]docleaner (only accessible from within the TUD campus network), which can be used to automatically remove potentially sensitive metadata from PDF documents. After processing a document, you will receive a brief overview of the original and subsequently removed metadata as well as a download option of the cleaned PDF. Your uploaded documents are kept on the server for a short period of time, after which they are automatically removed.

[1] https://www.zendas.de/themen/desktop/ms_office/verstecktedaten_2016.html
[2] https://www.zendas.de/login.html
[3] https://www.zendas.de/themen/desktop/ms_office/dokinspektor_2016.html
[4] https://support.microsoft.com/en-us/office/remove-hidden-data-and-personal-information-by-inspecting-documents-presentations-or-workbooks-356b7b5d-77af-44fe-a07f-9aa4d085966f
[5] https://helpx.adobe.com/acrobat/using/removing-sensitive-content-pdfs.html
[6] https://tools.pdf24.org/en/creator
[7] https://docleaner.cert.tu-dresden.de/