



## Toolbeitrag: Zotero

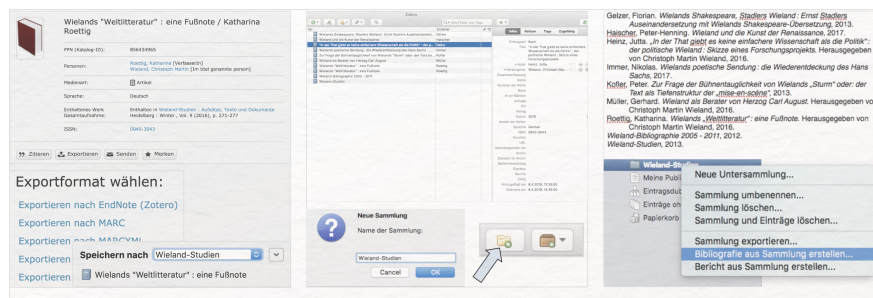
Marie Flüh  <sup>1</sup>

1. Universität Hamburg

forTEXT

Thema:	Bibliografie	DOI:	10.48694/fortext.3787
Jahrgang:	1	Ausgabe:	11
Erscheinungsdatum:	30-11-2024	Erstveröffentlichung:	2019-06-03 auf <a href="http://fortext.net">fortext.net</a>
Lizenz:			open  access

*Allgemeiner Hinweis: Rot dargestellte Begriffe werden im Glossar am Ende des Beitrags erläutert. Alle externen Links sind auch am Ende des Beitrags aufgeführt.*



*Der Zotero-Workflow: Nach dem Öffnen der grafischen Benutzeroberfläche lassen sich individuelle Sammlungen erstellen (s. Interface). Metadaten lassen sich entweder direkt in die Sammlung importieren (über den „Zotero-Connector“ oder die Exportfunktion der OPACs) oder händisch eintragen (s. Input). Schließlich können Sie Ihre Literaturlisten in unterschiedlichen automatisch angewandten Zitierformaten exportieren lassen (s. Output)*

- **Systemanforderungen:** Desktop-Version mit Windows-, Apple- und Linuxbetriebssystemen, Browsererweiterung (vgl. **Browser**) mit Firefox, Safari und Chrome nutzbar
- **Stand der Entwicklung:** Funktionalitäten der aktuellen Version (Zotero 5.0.66) werden stetig erweitert (Stand: 04.06.2019)
- **Herausgeber:** Zotero wird gegenwärtig von einer globalen Open Source Community weiterentwickelt und ist ursprünglich ein US-amerikanisches Kooperationsprojekt der Cooperation for Digital Scholarship und dem Roy Rosenzweig Center for History and New Media (George Mason University), gegründet von der Andrew W. Mellon Foundation, dem Institute of Museum and Library Services und der Alfred P. Sloan Foundation
- **Lizenz:** Kostenloses Open-Source (vgl. **Open Access**)-Tool, lizenzierte Erweiterungspakete mit erweiterten Funktionalitäten erhältlich
- **Weblink:** <https://www.zotero.org/>
- **Im- und Export:** Import (aus bibliographischen Datenbanken und anderen Literaturverwaltungsprogrammen) von 20 unterschiedlichen Dateiformaten, u. a. RIS, MODS, RDF, Bib TeX, Refer/BibIX, Import bibliographischer Daten auch durch Eingabe der ISBN, DOI, PubMed-ID, arXiv-ID; Export aus Datenbanken, die Austauschformate bieten (RIS, Bib TeX) oder Direktübernahme über Browser-Plugin „Zotero-Connect“, Export von Zotero-Sammlungen (als automatisch generierte Bibliographie als RTF (vgl. **Reintext-Version**), **HTML** oder als Sammlung in 17 verschiedenen Dateiformaten, u. a. BibLa Tex, RDF, JSON, **TEI**)
- **Sprachen:** Sprachunabhängig

## 1. Für welche Fragestellungen kann Zotero eingesetzt werden?

Mit der Open-Source-Software Zotero können Sie Forschungsliteratur sammeln, organisieren, teilen und bearbeiten. Zotero ermöglicht außerdem das persistente Zitieren sämtlicher Quellen in unterschiedlichen Zitierstilen und erleichtert durch eine große Bandbreite unterschiedlicher Funktionen und eine benutzerfreundliche Oberfläche (vgl. **GUI**) das Digitale Bibliografieren (Flüh 2024).

## 2. Welche Funktionalitäten bietet Zotero und wie zuverlässig ist das Tool?

*Funktionen (Auswahl):* Das Literaturverwaltungssystem beinhaltet eine Vielzahl an Funktionalitäten, die laufend aktualisiert werden und durch die Installation unterschiedlicher Plug-ins erweitert werden können.

- Import von Dateien über einen Importfilter (mehr als 20 Dateiformate)
- kostenloser Speicherplatz von 300 MB, weiterer Speicherplatz kann kostenpflichtig erworben werden
- ort- und zeitunabhängiger Zugriff von verschiedenen Rechnern aus, sofern Sie einen Zotero-Account angelegt haben (Zotero kann auch ohne diesen Account verwendet werden, dann entfällt diese Möglichkeit)
- Sprachauswahl: Für das Interface stehen 57 Sprachen zur Verfügung
- **Organisation** von Forschungsliteratur: automatische Erstellung von Indizes aus eingegebenen Daten, Verknüpfen von Datensätzen untereinander (über Funktion „zugehörige“), Einbinden von Dokumenten in die Datenbank und ortsunabhängiger Zugriff, Abruf von **Metadaten** aus PDF-Dokumenten (vgl. **PDF**)
- **Zitation** von Forschungsliteratur: Auswahl aus einer Vielzahl von Zitierstilen (11 bei Erstinstallation, über 9000 weitere Stile und von Nutzer\*innen modifizierte Stile über <https://www.zotero.org/styles12>), Quellenangaben per Drag & Drop in Word- oder Google Docs-Dokumente einfügen, Erstellen, Herunterladen und Teilen statischer und dynamischer Bibliographien, direkte Verknüpfung mit den gängigen Textverarbeitungsprogrammen
- **Kollaboratives Arbeiten**: Freigabe von Datensätzen für andere Nutzer\*innen, Social-Software-Funktionen wie Personensuche/Gruppensuche (Teilen von Bibliografien, gemeinsames Erstellen von Bibliografien), Nutzung über mobile Endgeräte (<https://www.zotero.org/support/mobile13>)

*Zuverlässigkeit*: Zotero bietet verlässliche Funktionalitäten, die das Sammeln, Organisieren, Bearbeiten, Teilen und Zitieren relevanter Fachliteratur umfassen. Als benutzerfreundlich erweist sich nicht nur die übersichtlich gestaltete dreigliedrige Oberfläche, die den Hover-Effekt für die Erläuterung sämtlicher Symbole sinnvoll einsetzt. Darüber hinaus fungiert die automatisch durchgeführte Synchronisierung der Webapplikation mit der Desktopvariante v. a. für Nutzer\*innen als hilfreiche Erweiterung, die an unterschiedlichen Rechnern arbeiten und durch die ständig aktualisierte Sammelschnittstelle ort- und zeitunabhängig auf Bibliografien und die hier hinterlegten Dokumente zugreifen können.

Anders als bei anderen Zitierprogrammen wie z.B. Citavi ist die Recherche in Datenbanken aus der Software heraus nicht möglich (nur ISBN-/DOI-Suche) und ein Dublettencheck ist lediglich manuell durchführbar.

### 3. Ist Zotero für DH-Einsteiger\*innen geeignet?

Checkliste	✓ / teilweise / –
Methodische Nähe zur traditionellen Literaturwissenschaft	✓
Grafische Benutzeroberfläche	✓
Intuitive Bedienbarkeit	teilweise
Leichter Einstieg	✓
Handbuch vorhanden	✓
Handbuch aktuell	✓
Tutorials vorhanden	✓
Erklärung von Fachbegriffen	✓
Gibt es eine gute Nutzerbetreuung?	✓

Die grundlegenden Funktionalitäten des Tools – wie das Anlegen von Bibliotheken, das Einfügen von Texten oder der Import von Metadaten – sind intuitiv bedienbar. Allerdings sind nicht alle der zahlreichen Funktionen sofort ersichtlich. Zahlreiche Tutorials und Videos erleichtern den Einstieg in die Arbeit mit Zotero. Im Zotero-Forum (<https://forums.zotero.org/14>) werden häufig gestellte Fragen diskutiert und hilfreiche Hinweise für eine eloquente Verwendung des Programms im Peer-to-Peer-Modus angeboten. Darüber hinaus können Sie den User-Support nutzen, um offene Fragen oder Anmerkungen einzubringen. Durch die stetig aktualisierte Homepage oder den Blog (<https://www.zotero.org/blog/15>) werden Sie über die neuesten Entwicklungen und Plug-ins informiert. Bei *ZotFile* (erweiterte Optionen für die Organisation von PDF-Dokumenten), dem *Zotero Storage Scanner* (verbesserte Auffindbarmachung von möglichen Verlinkungen und Duplikaten) oder dem *Zotero DOI Manager* (erweitertes DOI-Management) handelt es sich bspw. um optionale Erweiterungen der Zotero-Basisvariante.

### 4. Wie etabliert ist Zotero in den (Literatur-)Wissenschaften?

Zotero hat sich in den Literaturwissenschaften neben kommerziellen Literaturverwaltungsprogrammen wie Citavi oder EndNote als ebenso performante wie benutzerfreundliche Alternative bewährt und etabliert.

## 5. Unterstützt Zotero kollaboratives Arbeiten?

Ja, mit Zotero ist die Kooperation in Arbeitsgruppen – deren Teilnehmer\*innenzahl und Privatsphäre-Einstellungen je nach verwendeter Zotero-Variante variiert – möglich. Auf diese Weise können Sie gemeinsam Bibliografien erstellen und bearbeiten. Im Online-Interface können Sie bereits erarbeitete Bibliotheken, Gruppen oder Paper auch für externe Nutzer\*innen als read-only-Variante freischalten.

## 6. Sind meine Daten bei Zotero sicher?

Ja, Zotero ist eine unabhängige nonprofit Organisation, deren Datenschutzrichtlinie die Weiterverwendung Ihrer persönlichen Daten ausschließt. Die Entscheidung, Datensätze zu teilen oder diese ausschließlich privat zu verwenden, liegt bei Ihnen. Sofern Sie sich keinen Online-Account anlegen und auf die damit einhergehenden erweiterten Funktionalitäten verzichten, werden Daten standardmäßig ausschließlich auf Ihrem eigenen Rechner gespeichert.

Datensicherheit ist vor allem bei der Verwendung der kostenlosen Desktop-Variante gewährleistet. Da die Speicherkapazität in der Grundversion begrenzt ist, kann zusätzlicher Speicherplatz käuflich erworben werden. Spätestens an dieser Stelle nutzen Sie eines der erweiterten Angebote von Zotero, womit andere Datenschutzrichtlinien geltend werden. Diese werden auch dann bereits wirksam, wenn Sie neben der kostenlosen Desktop-Variante die Synchronisierungs-Funktion Ihrer Daten nutzen. Um diese Funktion nutzen zu können, müssen Sie einen Zotero-Account mit einem individuellen Nutzer\*innenprofil erstellen. Gesammelt werden dann u. a.:

- personenbezogene Daten, die Sie bei der Erstellung eines Zotero-Accounts eingeben (Name und E-Mail-Adresse)
- Datensätze der Bibliografien, sofern Sie eine Synchronisierung mit dem Online-Account (Zotero-Server) umsetzen
- Daten angehängter Dateien (ebenfalls durch die Synchronisierung mit dem Zotero-Server (vgl. [Server](#)))
- die für den Zugang zu der synchronisierten Bibliothek verzeichnete IP-Adresse (vgl. [Ip-adresse](#))
- Informationen über Zugriffe auf die Zotero-Website (IP-Adresse, Browser)
- von Nutzer\*innen gesendete Fehlermeldungen

Um Nutzer\*innenanalysen durchführen zu können und im Zuge einer Weiterentwicklung des Tools werden anonymisierte Daten aggregiert, die bei der Veröffentlichung von Gruppenbibliotheken oder eigenen Publikationen entstehen (beide Funktionen sind ebenfalls nur via Online-Interface nutzbar). Die statistische Auswertung inkludiert ausschließlich frei zugängliche Metadaten (z. B. Titel der Publikation und Autor\*in).

Bei der Synchronisierung Ihrer Datensätze mit dem Zotero-Server werden Ihre Daten verschlüsselt und über die Amazon Cloud gesichert (us-east-1 AWS Region Virginia).

## Externe und weiterführende Links

- Zotero: <https://www.zotero.org/> (Letzter Zugriff: 06.11.2024)
- Zotero Blog: <https://web.archive.org/web/20241106113228/https://www.zotero.org/blog/> (Letzter Zugriff: 06.11.2024)
- Zotero Forum: <https://web.archive.org/web/20241106113309/https://forums.zotero.org/> (Letzter Zugriff: 06.11.2024)
- Zotero für mobile Endgeräte: <https://web.archive.org/web/20241106113344/https://www.zotero.org/support/mobile> (Letzter Zugriff: 06.11.2024)
- Zotero Zitierstile: <https://web.archive.org/web/20241106113414/https://www.zotero.org/styles> (Letzter Zugriff: 06.11.2024)

## Bibliographie

Flüh, Marie. 2024. Methodenbeitrag: Digitales Bibliografieren. Hg. von Evelyn Gius. *forTEXT* 1, Nr. 11. Bibliografie (29. November). doi: 10.48694/fortext.3786, <https://fortext.net/routinen/methoden/digitales-bibliografieren>.  
Puckett, Jason. 2011. *Zotero: A guide for librarians, researchers and educators*. Chicago: Assoc. of College and Research Libraries.

## Glossar

**Annotation** Annotation beschreibt die manuelle oder automatische Hinzufügung von Zusatzinformationen zu einem Text. Die manuelle Annotation wird händisch durchgeführt, während die (teil-)automatisierte Annotation durch **Machine-Learning-Verfahren** durchgeführt wird. Ein klassisches Beispiel ist das automatisierte **PoS-Tagging** (Part-of-Speech-Tagging), welches oftmals als Grundlage (**Preprocessing**) für

weitere Analysen wie Named Entity Recognition (NER) nötig ist. Annotationen können zudem deskriptiv oder analytisch sein.

- Browser** Mit Browser ist in der Regel ein Webbrowser gemeint, also ein Computerprogramm, mit dem das Anschauen, Navigieren auf, und Interagieren mit Webseiten möglich wird. Am häufigsten genutzt werden dafür Chrome, Firefox, Safari oder der Internet Explorer.
- Commandline** Die Commandline (engl. *command line interface* (CLI)), auch Kommandozeile, Konsole, Terminal oder Eingabeaufforderung genannt, ist die direkteste Methode zur Interaktion eines Menschen mit einem Computer. Programme ohne eine grafische Benutzeroberfläche (GUI) werden i. d. R. durch Texteingabe in die Commandline gesteuert. Um die Commandline zu öffnen, klicken Sie auf Ihrem Mac „cmd“ + „space“, geben „Terminal“ ein und doppelklicken auf das Suchergebnis. Bei Windows klicken Sie die Windowstaste + „R“, geben „cmd.exe“ ein und klicken Enter.
- CSV** CSV ist die englische Abkürzung für *Comma Separated Values*. Es handelt sich um ein Dateiformat zur einheitlichen Darstellung und Speicherung von einfach strukturierten Daten mit dem Kürzel `.csv`, sodass diese problemlos zwischen IT-Systemen ausgetauscht werden können. Dabei sind alle Daten zeilenweise angeordnet. Alle Zeilen wiederum sind in einzelne Datenfelder aufgeteilt, welche durch Trennzeichen wie Semikola oder Kommata getrennt werden können. In Programmen wie Excel können solche Textdateien als Tabelle angezeigt werden.
- GUI** GUI steht für *Graphical User Interface* und bezeichnet eine grafische Benutzeroberfläche. Ein GUI ermöglicht es, Tools mithilfe von grafischen Schaltflächen zu bedienen, um somit beispielsweise den Umgang mit der **Commandline** zu umgehen.
- HTML** HTML steht für *Hypertext Markup Language* und ist eine textbasierte Auszeichnungssprache zur Strukturierung elektronischer Dokumente. HTML-Dokumente werden von **Webbrowsern** dargestellt und geben die Struktur und Online-Darstellung eines Textes vor. HTML-Dateien können außerdem zusätzliche **Metainformationen** enthalten, die auf einer Webseite selbst nicht ersichtlich sind.
- Ip-adresse** Die Vernetzung von Computern wird in einem Internetprotokoll (IP) festgehalten, woraufhin jedes angebundene Gerät in diesem Computernetz eine IP-Adresse erhält. So werden die Geräte adressierbar und erreichbar gemacht. Die IP gehört zu den personenbezogenen Daten, da über sie auf Ihre Identität geschlossen werden kann.
- Lemmatisieren** Die Lemmatisierung von Textdaten gehört zu den wichtigen **Preprocessing**-Schritten in der Textverarbeitung. Dabei werden alle Wörter (**Token**) eines Textes auf ihre Grundform zurückgeführt. So werden beispielsweise Flexionsformen wie „schneller“ und „schnelle“ dem Lemma „schnell“ zugeordnet.
- Machine Learning** Machine Learning, bzw. maschinelles Lernen im Deutschen, ist ein Teilbereich der künstlichen Intelligenz. Auf Grundlage möglichst vieler (Text-)Daten erkennt und erlernt ein Computer die häufig sehr komplexen Muster und Gesetzmäßigkeiten bestimmter Phänomene. Daraufhin können die aus den Daten gewonnen Erkenntnisse verallgemeinert werden und für neue Problemlösungen oder für die Analyse von bisher unbekanntem Daten verwendet werden.
- Markup (Textauszeichnung)** Die Textauszeichnung (eng. *Markup*) fällt in den Bereich der Daten- bzw. Textverarbeitung, genauer in das Gebiet der Textformatierung, welche durch **Auszeichnungssprachen** wie XML implementiert wird. Dabei geht es um die Beschreibung, wie einzelne Elemente eines Textes beispielsweise auf Webseiten grafisch dargestellt werden sollen.
- Markup Language** Markup Language bezeichnet eine maschinenlesbare Auszeichnungssprache, wie z. B. **HTML**, zur Formatierung und Gliederung von Texten und anderen Daten. So werden beispielsweise auch **Annotationen** durch ihre Digitalisierung oder ihre digitale Erstellung zu Markup, indem sie den Inhalt eines Dokumentes strukturieren.
- Metadaten** Metadaten oder Metainformationen sind strukturierte Daten, die andere Daten beschreiben. Dabei kann zwischen administrativen (z. B. Zugriffsrechte, Lizenzierung), deskriptiven (z. B. Textsorte), strukturellen (z. B. Absätze oder Kapitel eines Textes) und technischen (z. B. digitale Auflösung, Material) Metadaten unterschieden werden. Auch **Annotationen** bzw. **Markup** sind Metadaten, da sie Daten/Informationen sind, die den eigentlichen Textdaten hinzugefügt werden und Informationen über die Merkmale der beschriebenen Daten liefern.
- Named Entities** Eine Named Entity (NE) ist eine Entität, oft ein Eigenname, die meist in Form einer Nominalphrase zu identifizieren ist. Named Entities können beispielsweise Personen wie „Nils Holgerson“, Organisationen wie „WHO“ oder Orte wie „New York“ sein. Named Entities können durch das Verfahren der Named Entity Recognition (NER) automatisiert ermittelt werden.
- OCR** OCR steht für *Optical Character Recognition* und bezeichnet die automatische Texterkennung von gedruckten Texten, d. h. ein Computer „liest“ ein eingescanntes Dokument, erkennt und erfasst den Text darin und generiert daraufhin eine elektronische Version.
- Open Access** Open Access bezeichnet den freien Zugang zu wissenschaftlicher Literatur und anderen Materialien im Internet.

- PDF** PDF steht für *Portable Document Format*. Es handelt sich um ein plattformunabhängiges Dateiformat, dessen Inhalt auf jedem Gerät und in jedem Programm originalgetreu wiedergegeben wird. PDF-Dateien können Bilddateien (z. B. Scans von Texten) oder computerlesbarer Text sein. Ein lesbares PDF ist entweder ein **OCR**ter Scan oder ein am Computer erstellter Text.
- POS** PoS steht für *Part of Speech*, oder „Wortart“ auf Deutsch. Das PoS- **Tagging** beschreibt die (automatische) Erfassung und Kennzeichnung von Wortarten in einem Text und ist of ein wichtiger **Preprocessing**-Schritt, beispielsweise für die Analyse von **Named Entities**.
- Preprocessing** Für viele digitale Methoden müssen die zu analysierenden Texte vorab „bereinigt“ oder „vorbe-reitet“ werden. Für statistische Zwecke werden Texte bspw. häufig in gleich große Segmente unterteilt (*chunking*), Großbuchstaben werden in Kleinbuchstaben verwandelt oder Wörter werden **lemmatisiert**.
- Reintext-Version** Die Reintext-Version ist die Version eines digitalen Textes oder einer Tabelle, in der keinerlei Formatierungen (Kursivierung, Metadatenauszeichnung etc.) enthalten sind. Reintext-Formate sind beispielsweise TXT, RTF und **CSV**.
- Server** Ein Server kann sowohl hard- als auch softwarebasiert sein. Ein hardwarebasierter Server ist ein Computer, der in ein Rechnernetz eingebunden ist und der so Ressourcen über ein Netzwerk zur Verfügung stellt. Ein softwarebasierter Server hingegen ist ein Programm, das einen spezifischen Service bietet, welcher von anderen Programmen (Clients) lokal oder über ein Netzwerk in Anspruch genommen wird.
- TEI** Die *Text Encoding Initiative* (TEI) ist ein Konsortium, das gemeinsam einen Standard für die Darstellung von Texten in digitaler Form entwickelt. Die TEI bietet beispielsweise Standards zur Kodierung von gedruckten Werken und zur Auszeichnung von sprachlichen Informationen in maschinenlesbaren Texten (siehe auch **XML** und **Markup**).
- Type/Token** Das Begriffspaar „Type/Token“ wird grundsätzlich zur Unterscheidung von einzelnen Vorkommnissen (Token) und Typen (Types) von Wörtern oder Äußerungen in Texten genutzt. Ein Token ist also ein konkretes Exemplar eines bestimmten Typs, während ein Typ eine im Prinzip unbegrenzte Menge von Exemplaren (Token) umfasst.  
Es gibt allerdings etwas divergierende Definitionen zur Type-Token-Unterscheidung. Eine präzise Definition ist daher immer erstrebenswert. Der Satz „Ein Bär ist ein Bär.“ beinhaltet beispielsweise fünf Worttoken („Ein“, „Bär“, „ist“, „ein“, „Bär“) und drei Types, nämlich: „ein“, „Bär“, „ist“. Allerdings könnten auch vier Types, „Ein“, „ein“, „Bär“ und „ist“, als solche identifiziert werden, wenn Großbuchstaben beachtet werden.
- XML** XML steht für *Extensible Markup Language* und ist eine Form von **Markup Language**, die sowohl computer- als auch menschenlesbar und hochgradig anpassbar ist. Dabei werden Textdateien hierarchisch strukturiert dargestellt und Zusatzinformationen i. d. R. in einer anderen Farbe als der eigentliche (schwarz gedruckte) Text dargestellt. Eine standardisierte Form von XML ist das **TEI-XML**.