

DANTE
Deutschsprachige
Anwendervereinigung T_EX e.V.

Jürgen Fenn: *Literaturverwaltung mit Firefox, Zotero und BIBT_EX*, Die T_EXnische Komödie 4/2007, S. 20–26.

Reproduktion oder Nutzung dieses Beitrags durch konventionelle, elektronische oder beliebige andere Verfahren ist nur im nicht-kommerziellen Rahmen gestattet. Verwendungen in größerem Umfang bitte zur Information bei DANTE e.V. melden. Für kommerzielle Nutzung ist die Zustimmung der Autoren einzuholen.

Die T_EXnische Komödie ist die Mitgliedszeitschrift von DANTE, Deutschsprachige Anwendervereinigung T_EX e.V. Einzelne Hefte können von Mitgliedern bei der Geschäftsstelle von DANTE, Deutschsprachige Anwendervereinigung T_EX e.V. erworben werden. Mitglieder erhalten Die T_EXnische Komödie im Rahmen ihrer Mitgliedschaft.

Literaturverwaltung mit Firefox, Zotero und BIBT_EX

Jürgen Fenn

Der Beitrag stellt eine neuentwickelte Literaturverwaltung namens **Zotero** vor, die als Plugin in **Firefox 2** eingebunden wird. Damit können auch BIBT_EX-Daten verarbeitet werden.

Das Web 2.0 entlässt seine Kinder

Der Hype um das „Web 2.0“ in den einschlägigen Foren, Blogs und Zeitschriften ist gerade etwas am Abklingen, da tauchen die ersten ernsthafteren Anwendungen auf, die – abseits von den mittlerweile allgegenwärtigen RSS-Feeds, Widgets und Podcasts – auch dem technisch weniger versierten Nutzer einen guten Eindruck davon vermitteln, was für ein Potential tatsächlich im „semantischen Web“ [9] stecken könnte und insbesondere wie sich die Technik des wissenschaftlichen Arbeitens durch ihren Einsatz schon sehr bald nachhaltig ändern wird.

Vor allem die Arbeit mit Online-Bibliographien und BIBT_EX [3] kann mithilfe des Add-On **Zotero** [12] zu dem Browser **Firefox 2** ganz wesentlich erleich-

tert werden. Aber auch die Arbeit mit anderen online verfügbaren Quellen in wissenschaftlichen Arbeiten vereinfacht sich damit teilweise erheblich.

Die Funktionsweise

Im Gegensatz zu einem grafischen Frontend wie **Jabref** [5], das BibTeX-Dateien unter einer grafischen Oberfläche verwaltet, kann man mit **Zotero** ganze Projekte organisieren. Hiermit können die zu zitierenden Quellen selbst ebenso wie die ihnen zugehörigen Metadaten systematisch archiviert, katalogisiert, kommentiert und untereinander verknüpft werden.

Die *Wikipedia* [10] zitiert zwei Meldungen, wonach die nächsten Versionen der Browser **Firefox** und **Internet Explorer** sogenannte *Mikroformate* nativ unterstützen werden. Web-Seiten können mit dieser Technik, vom Benutzer unbemerkt, mit Informationen über die Bedeutung des Texts, der Bilder oder anderer Objekte, die in ihnen enthalten sind oder die über sie abgerufen werden können, versehen werden. Diese zusätzlichen (Meta-)Informationen sind im HTML/XHTML-Quelltext der Seite enthalten. Hierfür können etwa die HTML-Tags `div` und `span` sowie das `class`-Attribut verwendet werden (vgl. hierzu bereits [4, Abschn. 7.4.1.4]). Diese Informationen können deshalb vom Browser, ebenso wie der übrige Inhalt einer Web-Seite, geparkt und automatisch weiterverarbeitet werden.

Für unser Thema sind dabei natürlich vor allem diejenigen Anwendungen von Interesse, die bibliographische Daten bereitstellen. Insbesondere das COinS-Format („*ContextObjects in Spans*“) hat sich mittlerweile etabliert, es wird beispielsweise auch in der *Wikipedia* verwendet. [11] Mithilfe von Mikroformaten stehen beim Aufruf einer Web-Seite im Browser sofort weitere Informationen zur Verfügung, die dann auch in eine Literaturverwaltung übernommen werden können. Ein übergreifender Standard namens „*cite*“ zur Auszeichnung von Zitaten befindet sich derzeit in der Diskussion. [6]

Hier setzt **Zotero** als Erweiterung von **Firefox 2** an, denn es wird als Plug-In betrieben und arbeitet deshalb nahtlos mit dem Browser zusammen, kann also direkt auf die Inhalte zugreifen, die im Browser beim Surfen im Netz ankommen. **Zotero** ist eine Literaturverwaltung im Browser, die alle möglichen online verfügbare Inhalte sammeln und später weiterverarbeiten kann. Sie ist auf jeder Plattform verfügbar, auf der **Firefox** läuft, also praktisch auf jedem heute gebräuchlichen Computer – ebenso wie *TeX & Friends*.

Eine Anwendung wie diese ist natürlich auf die Unterstützung der Anbieter im WWW angewiesen. Die Zahl der Seiten, die Mikroformate verbreiten, hat sich in der letzten Zeit schon beträchtlich vergrößert. Einen Überblick einschlägiger Seiten findet man unter <http://www.zotero.org/translators/>. Hier finden sich nicht nur „große Namen“ wie *Amazon*, [arXiv.org](http://arxiv.org), die Zeitschrift *Nature*, *Google Books* und *Google Scholar* sowie *CiteSeer* und die *Wikipedia* [11], sondern auch viele amerikanische Bibliothekskataloge, vor allem natürlich die *Library of Congress*. Der *Gemeinsame Bibliotheksverbund GBV* unter <http://www.gbv.de/vgm/> ist der erste deutsche Verbundkatalog, der mit *Zotero* zusammenarbeitet.

Mit Zotero arbeiten

Die Installation von *Zotero* ist denkbar einfach, sie erfolgt wie bei anderen Plugins auch automatisch beim Herunterladen bzw. beim Öffnen der *xpi*-Datei mit *Firefox* (Download-Link auf der Startseite der Projekt-Homepage <http://www.zotero.org/>). Das Plug-In steht nach einem Neustart des Browsers zur Verfügung, was an dem *Zotero*-Logo unten rechts in der Statuszeile des Browsers erkennbar ist. Wie bei anderen Plugins auch, ist ein automatisches Update möglich, wenn eine neue Version bereitsteht.

Klickt man auf das Logo, öffnet sich die grafische Benutzeroberfläche im unteren Teil des Browserfensters. Sie ist in drei Spalten geteilt, welche – von links nach rechts – sämtliche Quellen, organisiert in Ordnern (und Unterordnern); die einzelnen Quellen in dem jeweils markierten Ordner der Sammlung; sowie die bibliographischen Angaben eines im mittleren Fenster markierten Objekts zeigen (vgl. Abb. 1).

Per Kontextmenü können Titel ordnerweise oder einzeln in eine Datei zur weiteren Verarbeitung exportiert werden. Unterstützt werden neben *B_IB_TE_X* derzeit die Formate *MODS*, *Endnote*, *RIS*, *Unqualified Dublin Core XML* und die *Wikipedia*-Zitiervorlage.

Zotero speichert ganze Web-Seiten als *snapshot* komplett ab, so dass sie später lokal wieder angezeigt werden können. Jede Quelle kann – anscheinend ohne Begrenzung – mit beliebigen Notizen, Such-Tags oder Anhängen versehen werden. Es steht eine sehr große Zahl an Quellentypen zur Verfügung, von Büchern und Aufsätzen über Blog Postings bis hin zu Gerichtsentscheidungen, die bekanntlich zunehmend online im Volltext verbreitet und zitiert werden. Natürlich kann diese Datenbank durchsucht werden. Die gesamte

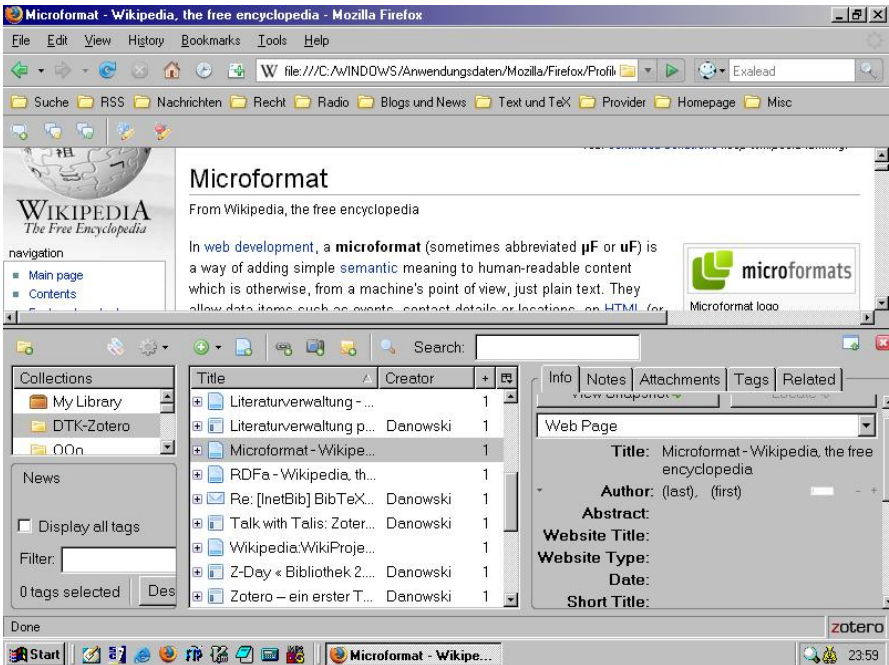


Abbildung 1: Zotero-Oberfläche bei der Darstellung des gespeicherten *snapshot* einer HTML-Seite

collection kann in dem eigenen Format „Zotero RDF“ mit allen Anhängen und Notizen gespeichert werden, um ein Backup zu erstellen.

Erfahrungen und Probleme

Obwohl sich das Programm derzeit noch im Beta-Stadium befindet, ist es schon ziemlich leistungsfähig. *Zotero* v1.0.0b4.r5 vom August 2007 lief im Test stabil.

Was für die Arbeit mit Online-Bibliographien allgemein gilt, [1, S. 91] ist auch hier zu beachten: Blind kann man dem Output nicht vertrauen. Fast alle Zitate des vorliegenden Artikels wurden mit *Zotero* erzeugt und im *Emacs* weiterbearbeitet. Alle BiBTeX-Datensätze, die aus *Zotero* exportiert wurden, waren formal valide, aber inhaltlich unvollständig, so dass sie händisch ergänzt werden mussten. Zitierte URLs mussten z. B. von dem `url`-Feld in das `note`-Feld des Eintrags verlegt werden, weil der Bibliographiestil der DTK das Feld `url` nicht unterstützt. Bei den Artikeln aus der *Wikipedia* ist es derzeit sogar einfacher, deren eigene Zitierfunktion [3, S. 44 f] zu verwenden.

Sucht man beispielsweise nach dem „L^AT_EX-Begleiter“ [7] im GBV, so wird in der Adresszeile des Browsers rechts ein Dokumentensymbol angezeigt. Klickt man hierauf (wie beim Abonnieren eines RSS-Feeds), speichert *Zotero* die bibliographischen Angaben zu dem Buch. Exportiert man sie ins BiBTeX-Format, erhält man einen Eintrag, der nach einer leichten Nachbearbeitung im *Emacs* wegen des Zeilenumbruchs wie in Abbildung 2 aussieht. Hier fehlen nicht nur einige Autorennamen (nämlich David Carlisle, Chris Rowley und Joachim Schrod), sondern auch die Angabe der Auflage, die im Katalog durchaus vorhanden ist. Auch der Schriftzug „L^AT_EX“ wurde nicht korrekt erkannt und umgesetzt.

Der Import bestehender BiBTeX-Dateien in *Zotero* funktioniert meist reibungslos. Kommt es hier zu Problemen, was auch bei zweifelsfrei validen `bib`-Dateien passieren kann – im Test war eine sehr große Datei betroffen –, ist es allerdings kaum möglich, den Fehler einzugrenzen, weil keine Fehlermeldungen erfolgen. Probleme ergeben sich vor allem bei der Weiterverarbeitung, wenn die Einträge L^AT_EX-Makros enthalten. Beim Import werden die geschweiften Klammern entfernt, der Backslash vor dem Makronamen wird aber übernommen. Exportiert man einen solchen Eintrag wiederum aus *Zotero* nach BiBTeX, so wird ein zweiter Backslash hinzugefügt, den man händisch entfernen muss. Auch der Export des maskierten scharfen S funk-

```

@book{mittelbach_der_2005,
  address =    {M{"\u}nchen [u.a.]},
  series =    {Scientific tools ; 7166},
  title =    {Der LATEX-Begleiter},
  isbn =     {3-8273-7166-X},
  publisher = {Pearson Studium},
  author =   {Frank Mittelbach and Michel Goossens and
             Johannes Braams and Christine Detig},
  year =    2005,
  keywords = {LATEX <Programm> ,LATEX 2E }
}

```

Abbildung 2: Eintrag des $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ -Begleiters [7] aus dem GBV, wie er von Zotero ins $\text{BIB}_{\text{T}}\text{E}_{\text{X}}$ -Format exportiert wurde

tioniert derzeit nicht korrekt. Umlaute werden dagegen richtig behandelt. Anwender berichten auch darüber, dass manchmal Autorennamen nicht vollständig exportiert würden, vielleicht ähnlich wie zuvor in Abb. 2 erkennbar.

Zusammenfassung und Ausblick

Zotero ist schon heute sehr hilfreich bei der Generierung von $\text{BIB}_{\text{T}}\text{E}_{\text{X}}$ -Einträgen für die eigene Bibliographie und sollte deshalb unbedingt einbezogen werden, wenn ein größeres Projekt mit umfangreicher Quellenarbeit bevorsteht. Die Entwicklung von Zotero verläuft sehr zügig, wie man in dem projekteigenen Blog unter <http://www.zotero.org/blog/> und in dem Blog von *Patrick Danowski* [2] verfolgen kann. Deshalb ist anzunehmen, dass die derzeit noch verbliebenen Kinderkrankheiten bald behoben sein werden und Zotero ein wichtiges und hilfreiches Tool für die Arbeit mit $\text{BIB}_{\text{T}}\text{E}_{\text{X}}$ wird, vor allem weil die Unterstützung für Mikroformate in den Bibliothekskatalogen auch hierzulande bald noch weiter zunehmen wird. Updates können über einen Schalter in Firefox unter Tools|Add-ons automatisiert werden.

Verglichen mit den Lösungen, die in [3] vorgestellt worden waren, zeigt das Programm noch einige Schwächen, ist aber ganz sicherlich auf einem guten Weg.

Wer sich für die weitere Entwicklung bei Werkzeugen zur Literaturverwaltung interessiert, sollte übrigens die Blogs und die Mailinglisten für Bibliothekare, etwa das gerade genannte Blog von *Danowski* [2], verfolgen. Bekannt ist

insbesondere die von *Michael Scharwächter* betriebene *INETBIB*-Mailingliste [8], die auch über ein durchsuchbares Archiv verfügt. Man kann die Liste auch über den Newsserver <news://news.gmane.org> lesen, wo sie unter gmane.culture.libraries.inetbib geführt wird.

Literatur

- [1] Nelson H. F. Beebe: *A Bibliographer's Toolbox*; *TUGboat*; 25(1), S. 89–104; 2004; Proceedings of the Practical T_EX 2004 Conference, <http://www.tug.org/TUGboat/Articles/tb25-1/beebe-bib.pdf>.
- [2] Patrick Danowski: *Bibliothek 2.0 und mehr*; <http://www.bibliothek2null.de/?p=187>, früher unter: <http://bibliothek2.wordpress.com/>.
- [3] Jürgen Fenn: *Online-Bibliographien nutzen mit BibT_EX*; *Die T_EXnische Komödie*; 4/06, S. 40–46; 2006.
- [4] Michel Goossens, Sebastian Rahtz et al.: *The L^AT_EX Web Companion. Integrating T_EX, HTML, and XML*; Tools and Techniques for Computer Typesetting; Addison-Wesley; Boston et. al.; 1999; eighth printing, January 2006.
- [5] JabRef Projekt-Homepage: <http://jabref.sourceforge.net/>.
- [6] Microformats Wiki: *citation – Microformats*; 2007; <http://microformats.org/wiki/cite>, Zugriff am 13. August 2007.
- [7] Frank Mittelbach, Michel Goossens et al.: *Der L^AT_EX-Begleiter*; Scientific tools ; 7166; Pearson Studium; München [u.a.]; 2. Aufl.; 2005.
- [8] Michael Scharwächter: *INETBIB-Mailingliste. Internet in Bibliotheken*; <http://inetbib.de/>.
- [9] W3C: *Semantic Web Activity*; <http://www.w3.org/2001/sw/>.
- [10] Wikipedia: *Microformat – Wikipedia, the free encyclopedia*; 2007; <http://en.wikipedia.org/wiki/Microformat>, Zugriff am 13. August 2007.
- [11] Wikipedia: *Wikipedia:WikiProject Microformats/COinS – Wikipedia, the free encyclopedia*; 2007; http://en.wikipedia.org/wiki/Wikipedia:WikiProject_Microformats/COinS, Zugriff am 13. August 2007.

- [12] Zotero Projekt-Homepage: *Zotero - The Next-Generation Research Tool*; 2007; <http://www.zotero.org/>.