

CTAN 2025: Architektur, Betrieb, Team und Zukunft

Erik Braun

CTAN ist die zentrale Infrastruktur für \TeX -Pakete und -Software. Dieser Beitrag skizziert Ursprung, Team, Architektur, Spiegelnetz, Upload-Prozess, Kennzahlen sowie zukünftige Entwicklungen. Inhaltlich ist er am Vortrag »The state of CTAN: 2025« [1], der auf der Konferenz der \TeX Users Group (TUG) 2025 in Trivandrum/Indien gehalten wurde, angelehnt.

Die Anfänge

Das *Comprehensive \TeX Archive Network* (CTAN) entstand Anfang der 1990er Jahre als Antwort auf die fragmentierte Verteilung von \TeX -Paketen über Mailinglisten, FTP-Server und Usenet-Newsgroups. Laut Wikipedia [6] ging die Initiative zur Entwicklung von CTAN von Joachim Schrod aus, der eine Reihe von namhaften Mitstreitern gewinnen konnte, die ein verteiltes Netzwerk zum Hochladen und Bereitstellen von \TeX -verwandter Software erschufen.

Eine ausführliche Konzeption findet sich in der Mitgliederzeitschrift der TUG »TUGboat« 3/1993 [2]. Im Großen und Ganzen wird CTAN wie vor über 30 Jahren geplant fortgeführt.

CTAN heute

CTAN beherbergt Pakete, Klassen, Fonts, Dokumentation und Werkzeuge rund um \TeX . Der *Catalogue* verknüpft installierbare Pakete, Metadaten, Kurzbeschreibungen und Verweise auf die Quellen.

Die Distributionen arbeiten eng mit CTAN zusammen, um die Daten in einem für sie komfortablen Format zu erhalten.

Der größte Unterschied zur Urfassung ist, daß statt mehrerer Hauptserver nur noch ein einziger verwendet wird. Zudem etablierten sich über 100 sogenannter Spiegel-Server, die zur regionalen Verteilung des Archivs dienen. CTAN stellt die Daten zu Verfügung, pflegt den umfangreichen strukturierten Katalog und stellt zudem die Verteilstation für schnelle Downloads dar.

Während in den frühen Jahren CTAN der Hauptanlaufpunkt für \TeX -Anwender war, können diese heutzutage darauf bauen, daß die \TeX -Distributionen aktuelle Versionen der \TeX -Pakete bereitstellen. Diese Distributionen, das sind führend TeXlive und MikTeX, beziehen ihrerseits ihre Daten direkt von CTAN.

Die \TeX nische Komödie ???

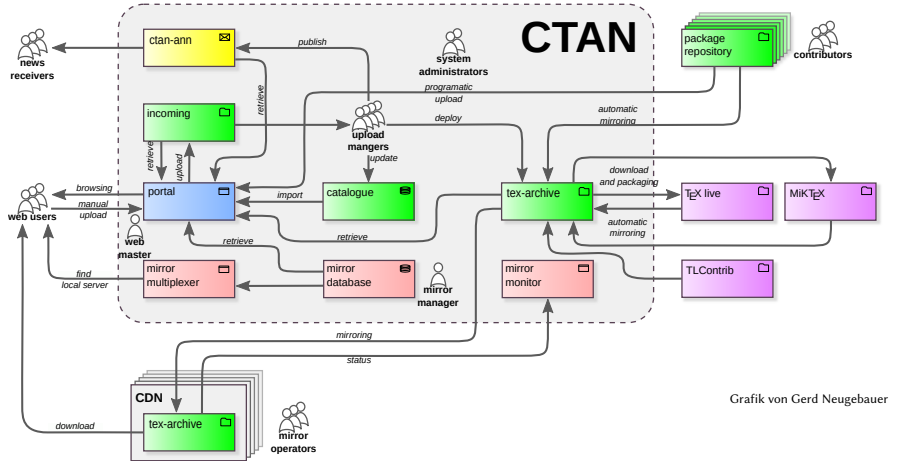


Abb. 1: Rollen und Datenfluß innerhalb von CTAN

Gerd Neugebauer hat das Zusammenspiel zwischen den \TeX -Paket-Entwicklern und Anwendern auf der einen Seite, dem CTAN-Team als Mittler, sowie den Spiegelserver-Betreibern und den Distributionen auf der anderen Seite in der anschaulichen Grafik 1 zusammengefaßt.

Team und Rollen

Im Sommer 2025 bestand das Team aus fünf Ehrenamtlichen: Petra Rube-Pugliese (seit 2014), Manfred Lotz (ebenfalls seit 2014), Erik Braun (seit 2016, als Reaktion auf [5]), Gerd Neugebauer (2000) und als jüngstem Mitglied Vincent Goulet (2025).

Das CTAN-Team sorgt nicht nur für die Infrastruktur und Uploads, sondern betreibt auch das Web-Portal samt Katalog. Für \TeX -Anwender wie auch die meisten Paketautoren ist genau diese Oberfläche die zentrale Anlaufstelle. Die Software dafür ist eine Eigenentwicklung von Gerd Neugebauer.

Das Upload-Management ist die beinahe tägliche Arbeit, die den größten Zeitaufwand benötigt. Bis auf Gerd, der als einziger Programmierer voll ausgelastet ist, bearbeiten als Team-Mitglieder die hochgeladenen Pakete: Sie schauen bei neuen Einreichungen auf die Formalien – also Lizenz, Versionsangaben usw. – und halten Katalog und Verzeichnisdaten aktuell. Außerdem sind sie in ständigem Kontakt mit den Autoren und mit den \TeX -Distributionen. Wenn ein neues Paket eingereicht wurde oder ein Autor das bei einem Update seines Pakets wünscht,

erfolgt eine offizielle Ankündigung über die Mailingliste `ctan-ann@ctan.org`, in der Usenet-Newsgruppe `comp.text.tex`¹ sowie auf dem CTAN-Portal.

Die Liste der Spiegelserver wird von Petra aktuell gehalten, sie kümmert sich ebenso um die Kommunikation mit den entsprechenden Serverbetreibern bei Problemen wie Nicht-Erreichbarkeit oder abgelaufenen TLS-Zertifikaten.

Weiterhin kümmert sich das Team um Betriebssysteminstallation und Dienste wie den Mailserver, sowie um die Erstellung von Programmen, die automatisiert die Konsistenz der Uploads und des Katalogs sicherstellen sollen.

Architektur und Plattform

CTAN ist so konzipiert, daß die einzelnen Aufgabenbereiche voneinander getrennt sind. Das erleichtert nicht nur Änderungen und Aktualisierungen, sondern eröffnet auch die Möglichkeit, Komponenten an engagierte Freiwillige zu übergeben, die sich gern einbringen möchten.²

Als Betriebssystem kommt mit Debian eine weit verbreitete und gut dokumentierte Linux-Distribution zum Einsatz. Das Portal selbst ist als Servlet-Anwendung in Tomcat realisiert, davor sitzt Apache als Handler. Für die Weiterleitung auf die Mirrors kommt die GeoIP-Implementierung der Firma MaxMind zum Einsatz.

Zur Infrastruktur gehören drei von CTAN gemietete Server. Der Portal-Server ist vor allem auf viel RAM und Plattenkapazität ausgelegt, um die wachsende Anzahl an Paketen bereitzustellen und das Portal im Hauptspeicher zu halten. Der Mirror-Master dagegen ist auf hohen Datendurchsatz optimiert; mit einer schnellen SSD und einer Netzwerkanbindung von 2,5 GBit/s dürften auf absehbare Zeit keine Engpässe bei der Synchronisierung der Mirrors auftreten.

Spiegelnetz und Synchronisation

Das CTAN-Netz umfasst weit über einhundert aktive Mirrors. Teilnehmer müssen mindestens einmal täglich mit dem Hauptserver von CTAN aktiv die Daten abgleichen. Die neue Netzwerkanbindung erlaubt deutlich häufigere Updates, so daß für Spiegelbetreiber nichts dagegen spricht, alle 6 Stunden oder auch stündlich eine Synchronisation durchzuführen.

Die Spiegelserver werden weltweit von Freiwilligen betrieben, mit einer deutlichen Häufung von Angeboten in Deutschland, China und Nordamerika (siehe Abbildung 2). In weniger abgedeckten Regionen sind TeX-Enthusiasten aufgerufen, zusätzliche

¹ Ja, es fällt einigen Leuten auf, wenn diese Ankündigungen nicht gepostet werden. Zu `comp.text.tex` siehe auch in der TUGboat 3/2024 [4].

² <mailto:ctan@ctan.org> 😊

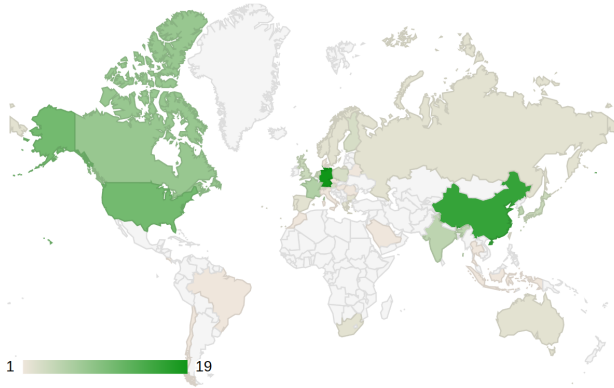


Abb. 2: Verbreitung der CTAN-Spiegel-Server

Mirrors bereitzustellen. Erforderlich sind im wesentlichen ein DNS-Name, mindestens 60 GB (besser 100 GB) Plattenplatz, ein Webserver-Dienst, eine dauerhafte Internetanbindung mit mindestens 1 Gbit/s sowie ein mindestens täglicher Synchronisationslauf (z. B. als rsync-Cronjob) zum Abgleich mit dem Hauptknoten. Weitere Informationen finden sich unter <https://ctan.org/mirrors/register/>.

Spiegelserver werden mithilfe des altherwürdigen, aber noch immer weitverbreiteten Programms »mirmon« regelmäßig daraufhin überprüft, wann sie das letzte Mal synchronisiert wurden. Zu beachten ist, daß (noch) keine Kontrolle dahingehend stattfindet, ob die gespiegelten Daten unverändert und vollständig vorliegen. Es wird das Änderungsdatum eines Zeitstempels im Wurzelverzeichnis betrachtet.

Der Zugriff auf die Spiegelserver erfolgt zufällig aus einer Liste der geografisch nahe liegenden Mirrors.³ Sollte die letzte Aktualisierung länger als 24 Stunden zurückliegen, so werden die betroffenen Spiegelserver automatisch temporär aus der Auswahlliste genommen.

Uploads und ihre Bearbeitung

Wie ein Paket auf CTAN landet, läuft nach einem festen Schema ab. Zunächst erfolgt die Einreichung über das Webformular auf <https://ctan.org/upload>. Alternativ kann der Upload auch automatisiert über eine Programmierschnittstelle erfolgen. Diese CTAN-API ist als stabil deklariert, bietet die Validierung sowie das Hochladen eines Pakets an und ist unter <https://ctan.org/help/submit> erläutert. Es müssen mindestens Paketname, Kurzbeschreibung, Version, Lizenzangaben

³ Das ist nicht optimal, da sich Netzwerkverbindungen nicht an Ländergrenzen halten. Freiwillige?

und ein Ansprechpartner angegeben werden, woraufhin das hochgeladene Paket manuell durch einen Upload-Manager bearbeitet wird. Welche Pakete noch nicht in Bearbeitung sind, ist öffentlich über <https://ctan.org/incoming> einsehbar. Erst nach erfolgreicher formaler Prüfung wird das Paket im passenden Verzeichnis des TeX-Archivs abgelegt. Zudem wird das Paket ins Katalogsystem übernommen – das heißt, geänderte Angaben werden in XML-Dateien hinterlegt und vorm Freigeben ins Portal syntaktisch validiert – und kann nun von den weltweiten Mirrors gespiegelt werden.

Um diesen Prozeß zu erleichtern, haben sich mehrere Hilfsprogramme etabliert:

- **13build** wird vom *L^AT_EX Project* bereitgestellt und dient dessen Mitarbeitern zur Entwicklung von hochwertigem L^AT_EX-Code. Dafür stehen zur Verfügung:
 - ein System für Unit-Tests
 - automatisiertes Setzen von Quelltexten
 - Erstellung der Upload-Pakete und Hochladen derselben

Ursprünglich wurde das Ganze zur Unterstützung der L^AT_EX-Entwicklung geschrieben, ist aber so angelegt, daß es auch problemlos von anderen Autoren genutzt werden kann.

- **pkgcheck** von Manfred Lotz ist *das* Standardwerkzeug für die lokale Vorprüfung. Es kontrolliert unter anderem Paketstruktur, Dateinamen, Zeilenenden, Zeichencodierung, Dateiberechtigungen und das Vorhandensein erforderlicher Dateien wie README. Auch TDS-Archive werden auf Konsistenz geprüft. Autoren wird empfohlen, es vor dem Einreichen selbst ausführen. Meldet pkgcheck keine Fehler, wird auch das CTAN-Team bei den geprüften formalen Aspekten nichts beanstanden.
- **ctanify** von Scott Pakin ist ein älteres, aber noch gebräuchliches Tool. Es erstellt aus den Quell- und Begleitdateien ein `.tar.gz`-Archiv in der von CTAN gewünschten Verzeichnisstruktur. Damit lassen sich Pakete ohne manuellen Aufwand CTAN-gerecht bündeln.

Diese Programme helfen dabei, Flüchtigkeitsfehler zu vermeiden und Pakete in konsistenter Struktur bereitzustellen. Wer regelmäßig auf CTAN veröffentlicht, spart damit nicht nur Zeit, sondern verringert auch das Risiko, daß ein Upload wegen formaler Mängel vom CTAN-Team zurückgewiesen wird.

Kennzahlen und Aktivität

Diese Zahlen beruhen auf dem Vortrag von Manfred auf der DANTE-Frühjahrstagung 2025. Die Folien dazu werden in [3] bereitgestellt. Die folgenden vermutlich nur für direkt Beteiligte interessanten Zahlen lassen sich in einem Satz zusammenfas-

Größe des T_EX-Archivs auf CTAN:

Stichtag	Archivgröße	Dateianzahl	Verzeichnisanzahl
2017-12-31	14G	285205	16353
2018-12-31	16G	301705	16887
2019-12-31	17G	329041	17261
2020-12-31	18G	337077	17336
2021-12-31	19G	357642	17628
2022-12-31	22G	363846	18171
2023-12-31	23G	374833	18447
2024-12-31	24G	387994	18910

(Hinweis: T_EXLive und MacT_EX mit insgesamt 32 GB sind nicht einbezogen.)

Anzahl der Autoren:

Jahr	Anzahl	davon ♀	Institutionen	verstorben
2017	2479	–	49	12
2018	2568	–	50	13
2019	2651	–	53	14
2020	2753	65	59	16
2021	2829	65	59	21
2022	2905	66	65	24
2023	2960	66	69	24
2024	3030	66	73	27

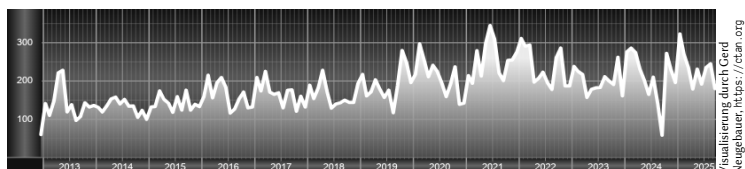
(Hinweis: Institutionen sind hauptsächlich Verleger und lokale Nutzergruppen.)

Anzahl der Pakete und Katalogeinträge auf CTAN:

	insgesamt	Daten auf CTAN	ohne Betreuer	Daten extern
2017-12-31	5433	5371	110	62
2018-12-31	5615	5555	114	60
2019-12-31	5778	5707	118	71
2020-12-31	5963	5882	134	81
2021-12-31	6161	6078	134	83
2022-12-31	6359	6275	138	84
2023-12-31	6528	6445	179	83
2024-12-31	6681	6597	197	84

(Hinweis: Nicht auf CTAN verfügbare Pakete sind Projekte wie Jabref, Ghostscript und pstotext.)

Anzahl der monatlichen Uploads auf CTAN:



sen: Die TeX-Gemeinschaft ist aktiv und zahlreich, und CTAN ist ein integraler Bestandteil.

Die Zahl der Pakete und Katalogeinträge nimmt moderat zu, im Katalog gibt es nur wenige Tage ohne neue Commits. Das aktive Datenvolumen umfaßt inzwischen etwa 25 GByte zuzüglich den Distributionen. Zusammengenommen erzeugen Portal und Mirror-Verkehr ein monatliches Datenvolumen im deutlich zweistelligen Terabyte-Bereich.

Dauerbaustellen

Theodore Roosevelt wird das Bonmot zugeschrieben, daß nur derjenige keine Fehler macht, der gar nichts unternimmt („*The only man who makes no mistakes is the man who never does anything.*“). Zum Glück ist die TeX-Gemeinschaft alles andere als untätig – was zwangsläufig dazu führt, daß bestimmte Probleme immer wieder auftreten. Wie bereits erwähnt, steht das CTAN-Team in engem Austausch mit den Distributionen, dem L^ATeX-Projekt und vor allem mit den Paketautoren. Um die Arbeit sowohl für Uploader als auch für das Team zu erleichtern, stehen zahlreiche Hilfestellungen und Anleitungen bereit.

Ein wiederkehrendes Thema sind **verwaiste Pakete**: Wenn der ursprüngliche Autor nicht mehr erreichbar ist und die Pflege stagniert, bleibt – auch abhängig von der Lizenz – die Wahl zwischen Maintainer-Wechsel, einer Kennzeichnung als veraltet oder einem *Fork*, also einer Abspaltung des Projekts durch die Community. Die am häufigsten verwendete *LaTeX Project Public License* erlaubt einen Wechsel des Maintainers, wenn der bisherige nicht mehr erreicht werden kann. Das ist auch der bevorzugte Weg, da auf diese Weise keine Dokumente wegen geänderter Paketnamen angepaßt werden müssen. Kann der Interessent den Autoren nicht erreichen, bemühen sich die CTAN-Mitarbeiter um Kontaktaufnahme; mitunter auch unter Verwendung von Kontaktdaten, die die Autoren nicht öffentlich angegeben, aber intern hinterlegt haben. Gelingt das über einen angemessenen Zeitraum hinweg nicht, wird das Paket als verwaist erklärt und kann von einem anderen Autor übernommen werden. Weitaus häufiger kommt es jedoch vor, daß der bisherige Autor selbst den neuen Maintainer benennt oder sich – durch das Interesse der TeX-Anwender angeregt – entschließt, seine Software nach längerer Pause wieder zu pflegen.

Auch lange etablierte Pakete können zu einer Herausforderung werden. Im Laufe der Jahre gewachsene Strukturen entsprechen häufig nicht mehr den heutigen Konventionen. Ein weiteres immer wieder auftretendes Thema sind seit langem vielbenutzte Pakete, die im Laufe der Zeit gewachsen sind und heute nicht mehr den üblichen Konventionen entsprechen („*legacy code*“). Einige Pakete verwenden

noch historisch akzeptierte Verzeichnislayouts oder enthalten implizite Abhängigkeiten, die unnötig komplex geraten sind. Zwar stehen inzwischen verschiedene Prüftools zur Verfügung, die viele dieser Probleme erkennen, doch für eine vollständige Umstruktur ist viel Handarbeit vonnöten, von CTAN und vom Betreuer des Pakets. Manche der gewachsenen Strukturen ließen sich nur mit so viel manuellem Aufwand an heutige Standards anpassen, daß man sie, wo sie keinen Schaden anrichten, einfach bestehen läßt.

Die Upload-Manager betrachten nur rein formale Probleme, oder anders gesagt: der Inhalt ist unerheblich, denn niemand kann alle Fachgebiete der Einreichungen im genügenden Maße kennen.

Zumindest ist das „eigentlich“ der Fall.

Ein Phänomen der letzten Jahre sind sehr einfache Pakete, die im Grunde nur einige persönliche Abkürzungen oder Makrosammlung ihrer Autoren enthalten. Häufig stammen sie von jungen oder unerfahrenen Anwendern, die gerade erst beginnen, sich mit \TeX und \LaTeX zu beschäftigen. Aus pädagogischer Sicht ist dies durchaus erfreulich, denn es zeigt, daß die Einstiegshürde niedriger als oft kommuniziert ist und Interesse an eigener Entwicklung besteht. Für die Allgemeinheit haben solche Pakete jedoch meist nur geringen Nutzen, da sie kaum über den persönlichen Arbeitsstil ihres Autors hinausreichen. Entsprechend stellt sich bei derartigen Einreichungen immer wieder die Frage, ob sie Teil des Archivs werden sollten oder besser in privater Nutzung verbleiben.

Ein neuer, bisher nur einmal aufgetretener Fall betraf ein Paket, das formal völlig korrekt aufgebaut war, dessen Inhalt sich jedoch bei näherem Hinsehen als inhaltlich sinnfrei erwies. Der Code war anscheinend mit Hilfe einer KI-Software erzeugt worden, und das Ergebnis bestand aus plausibel klingenden, aber letztlich bedeutungslosen Definitionen. Da die formalen Kriterien erfüllt waren und nicht völlig ausgeschlossen werden konnte, daß doch ein sinnvoller Kern dahintersteckte, wurde das Paket zunächst in das Archiv aufgenommen. Kurz darauf gab es ablehnende Reaktionen in Foren und auf Mailinglisten zu dieser Aufnahme auf CTAN (danke dafür). Der Vorfall zeigte, daß künftig auch die inhaltliche Plausibilität stärker berücksichtigt werden muß, selbst wenn Syntax und Lizenz formal einwandfrei sind.

Eine Vorgehensweise für diese Probleme wurde bisher nicht erarbeitet und wird sicher noch für Diskussionen sorgen.

Zukünftige Pläne für CTAN

Nach mehr als einem Jahrzehnt stabilen Betriebs steht das Portal www.ctan.org vor einem größeren technischen und gestalterischen Wandel. Die bisherige Plattform

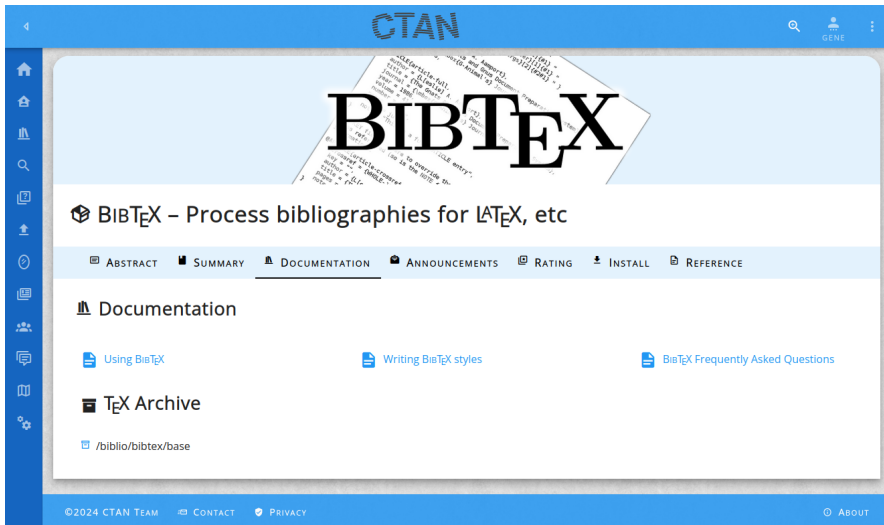


Abb. 3: Arbeitsbeispiel einer Paket-Seite (Dokumentation) der künftigen Portalsoftware

hat sich über viele Jahre hinweg bewährt, stößt jedoch zunehmend an ihre technischen Grenzen. Einige der zugrundeliegenden Komponenten sind veraltet oder werden nicht mehr aktiv gepflegt, was sowohl die Wartung als auch zukünftige Erweiterungen erschwert.

Die zukünftige Version des Portals, Codename »CTAN 3.0«, plant ist eine stärker clientseitige Darstellung unter Verwendung modernerer Webtechnologien wie Nuxt.js als Client-Framework und Dropwizard als serverseitige Anwendungsschicht.

Der unter der 3-Klausel-BSD-Lizenz veröffentlichte Quellcode dieser zum größten Teil von Gerd Neugebauer getragenen Neuentwicklung ist auf GitLab verfügbar, so daß sich jeder Interessiere an der Programmierarbeit beteiligen, Fehlerreports einreichen und Vorschläge unterbreiten kann. Über den Test- und Build-Server <https://build.ctan.org> findet sich nicht nur die Software und die Build- sowie Test-Umgebung, sondern auch ausführliche Dokumente zur Programmierung und der UI-Design-Guide. Dieser über 300 Seiten starke Leitfaden für Benutzeroberflächengestaltung beschreibt die gestalterischen und funktionalen Vorgaben für Aufbau, Aussehen und Verhalten der CTAN-Weboberfläche. Ein Beispiel, wie eine Paketseite ausshen könnte, findet sich in Abbildung 3.

Die Spiegelinfrastruktur bleibt von solchen Änderungen unberührt. Dieses System hat sich über viele Jahre als zuverlässig erwiesen, ist leicht zu betreiben, transparent dokumentiert und durch die große Zahl freiwillig gepflegter Mirrors für die meisten

Anwender gut zu erreichen. Eine Umstellung auf modernere oder zentralisierte Verteilmechanismen würde keinen nennenswerten Nutzen bringen, wohl aber zusätzliche Abhängigkeiten und Fehlermöglichkeiten schaffen. Vor allem aber würde sie das Risiko erhöhen, sich in technischer oder organisatorischer Hinsicht von kommerziellen Anbietern abhängig zu machen. Solange die bestehende Struktur stabil funktioniert und den Bedarf erfüllt, besteht daher kein Anlaß, an diesem bewährten Aufbau etwas zu ändern.

Fazit

CTAN ist und bleibt die tragende Infrastruktur des $\text{T}_{\text{E}}\text{X}$ -Ökosystems. Die Zahlen zeigen ein stetig wachsendes Archiv, stabile Prozesse und ein aktives Umfeld aus Autoren, Distributoren und freiwilligen Helfern.

Gleichzeitig ist der Portalumbau überfällig. Die geplante Modernisierung mit offenen Repositories und reproduzierbarer CI/CD-Pipeline überarbeitet die Schwachstellen der bisherigen Plattform.

Wie in jeder ehrenamtlichen Struktur und in jedem freien Softwareprojekt sind engagierte Mitwirkende ausdrücklich willkommen.

Weiterführendes

- [1] Erik Braun: The state of CTAN: 2025, Slides, TUG 2025, Accessed via TUG presentation archive, 2025, <https://tug.org/tug2025/av/d1-t02-braun-ctan/d1-t02-braun-ctan-slides.pdf>.
- [2] George D. Greenwade: »The Comprehensive TEX Archve Network (CTAN)«, *TUGboat*, 14.3 (1993), 342–351, <https://www.tug.org/TUGboat/Articles/tb14-3/tb40green.pdf>.
- [3] Manfred Lotz in Tagungsarchiv DANTE e.V. 2025, <https://www.dante.de/veranstaltungen/tagungsarchiv/tagungsarchiv-dante-2025/>.
- [4] Rayner Lucas, Tristan Miller, Marco Moock: »Usenet and the future of comp.text.tex«, *TUGboat*, 45.3 (2024), 299–301, DOI 10.47397/tb/45-3/tb141lucas-usenet, <https://www.tug.org/TUGboat/tb45-3/tb141lucas-usenet.pdf>.
- [5] Petra Rube-Pugliese: »CTAN sucht Mitstreiter«, *Die TeXnische Komödie*, 26.4 (2014), 64–67.
- [6] Wikipedia: CTAN (Comprehensive TeX Archive Network), <https://en.wikipedia.org/w/index.php?title=CTAN&oldid=1301529293>, [Online; accessed 2 Sep 2025], 2025.