

Gefördert durch:



Bundesministerium  
für Forschung, Technologie  
und Raumfahrt

**HTW****D**

Hochschule für Technik und  
Wirtschaft Dresden  
University of Applied Sciences



## PANTER – Publishing Analytics Tracker

<https://projekt-panter.de>

Sachbericht zum Verwendungsnachweis  
Teil I: Kurzbericht (Stand 12.02.2026)

ZE:	Förderkennzeichen:
Hochschule für Technik und Wirtschaft Dresden	16KOA022B
Vorhabenbezeichnung:	
<b>Verbundprojekt: Publishing Analytics Tracker – PANTER; Teilvorhaben B: Implementierung des Systems</b>	
Laufzeit des Vorhabens:	
<b>01.09.2023 - 31.08.2025</b>	
Ansprechpartner:	
<b>Prof. Dr.-Ing. Dietrich Kammer</b> Professur Technische Visualistik, Fakultät Informatik/Mathematik Hochschule für Technik und Wirtschaft Dresden (HTWD) Hochschule für angewandte Wissenschaften kammer@htw-dresden.de	

## Teil I: Kurzbericht

### I.1 Ursprüngliche Aufgabenstellung sowie wissenschaftlicher und technischer Stand

Ziel des Projektvorhabens **Publishing Analytics Tracker** (PANTER) war die Entwicklung eines zuverlässigen, öffentlich zur Verfügung stehenden Monitors für Finanzierungsmodelle, Preise und Preisentwicklungen auf dem Open-Access-Zeitschriftenmarkt. Das Projektvorhaben zielt damit auf eine der großen Informationslücken, die einer nachhaltigen und innovationsfreudigen Open-Access-Transformation im Wege stehen. Nur wenige Verlage weisen ihre Preise für Einreichungen in wissenschaftlichen Zeitschriften transparent aus. Damit werden Aussagen zur Preisentwicklung von Publikationsgebühren ebenso wie Preisvergleiche zwischen verschiedenen Zeitschriften unnötig erschwert bzw. technisch verhindert. Der Zugang zu Preisdaten ist in vielen Zeitschriften mit einem sehr hohen Aufwand verbunden und die Preismodelle so komplex, dass Autoren die Höhe der tatsächlichen Publikationsgebühren nicht ermitteln können.

Der PANTER-Preismonitor trägt durch die Veröffentlichung von Preisdaten und sonstigen relevanten Metadaten zu mehr Transparenz am akademischen Publikationsmarkt bei. Er ermöglicht den Vergleich verschiedener Zeitschriften und Verlage und gibt Auskunft über die Preisentwicklung der Publikationsgebühren. Damit bedient dieser den Bedarf verschiedener Nutzungsgruppen, die von Autoren über OA-Professionals bis hin zu Verhandlungsführern in den nationalen Konsortien reichen. Vergleichbare Projekte, die auf einer Datenbereitstellung durch Verlage aufbauen, sind in Bezug auf die Aktualität und Abdeckung der Daten vom Wohlwollen Dritter abhängig. Aus diesem Grund wurde in der Entwicklung von Beginn an auf einen Ansatz gesetzt, bei dem mittels Webscraping automatisiert Preisinformationen und sonstige Metadaten aus den Journalwebseiten extrahiert werden. Da im Projektverlauf auf der einen Seite die Grenzen statischer Verfahren des Webscrapings und auf der anderen Seite das Potential KI-basierter Verfahren immer deutlicher wurde, wird die Produktivpipeline durch große Sprachmodelle (LLM) unterstützt. Die validierten und in eine einheitliche Struktur gebrachten Daten aus dem Information-Retrieval-Prozess werden über eine REST-API in eine relationale Datenbank geschrieben. Ein interaktives Frontend erlaubt es den adressierten Nutzungsgruppen die benötigten Informationen zusammenzustellen, zu vergleichen und bei Bedarf per Export nachzunutzen. Das Projektvorhaben setzt auf den Aufbau einer modularen Informationsinfrastruktur und lässt sich für weitere Vorhaben/Services nachnutzen.

### I.2 Ablauf des Vorhabens

Das Projektvorhaben wurde von den drei Projektpartnern SLUB Dresden, der Hochschule für Technik und Wirtschaft Dresden (HTWD) sowie der Technischen Informationsbibliothek (TIB) beantragt und durchgeführt. Die Arbeitsaufgaben wurden entlang der Expertise der beteiligten Einrichtungen in drei Teilvorhaben mit den zugehörigen Arbeitspaketen (AP) untergliedert. Das Teilvorhaben I: Projektkoordination, Anforderungen, Datenerfassung und Datenmodellierung verantwortete der Projektpartner und Gesamtprojektleiter SLUB Dresden. Das Teilvorhaben II: Implementierung des Systems wurde von der HTWD und das Teilvorhaben III: Wissenschaftskommunikation und Vernetzung von der TIB übernommen. Die drei Teilvorhaben waren eng miteinander verzahnt und bauten in Teilen aufeinander auf.

Das Projekt startete im **AP 1** mit einer Anforderungsanalyse, in der im Rahmen eines Workshops zur nutzungsorientierten Gestaltung die relevanten Akteursgruppen in Form von Personas identifiziert und ihre Anforderungen an den Preismonitor bestimmt wurden (Grundlage Pflichtenheft). Ein weiterer Schwerpunkt des Arbeitspaketes bildete die Identifizierung der zentralen Publikationsorte für das deutsche Wissenschaftssystem und die Konzeption des Journal Price Index. Im **AP 2** wurde auf Basis der Anforderungsanalyse ein Datenbankschema erstellt, das alle für das Projektvorhaben relevanten Entitäten erfasst und über relationale Beziehungen verknüpft. Im **AP 3** wurde das Datenbankmodell im Datenbankmanagementsystem PostgreSQL implementiert und die Datenbank in Betrieb genommen. Das **AP 4** diente der Erfassung historischer Preisinformationen über das Internet Archive. Die bereits im AP 4 entwickelten Verfahren des Webscraping konnten im **AP 5** nachgenutzt und durch KI-basierte Verfahren in Richtung Produktivbetrieb erweitert werden. Im **AP 6** wurde unter Einbeziehung von Studierenden der Web-Service und das Frontend entwickelt. Die im **AP 7** entwickelte Schnittstelle bindet die Datenbank an das Frontend an. Im Rahmen des **AP 8** konnten wichtige Weichenstellungen in Richtung Communityaufbau/Vernetzung erreicht werden und es wurde die Qualität der Datenerhebungsverfahren sichergestellt. Das **AP 9** diente dazu, über Beiträge in Social Media, Fachvorträge und Workshops das zentrale Projektziel Preistransparenz zu kommunizieren. Der planmäßige Projektablauf sowie die Abstimmung der in Teilen ineinandergreifenden Arbeitspakete wurde über das **AP 10** koordiniert.

### **I.3 Wesentliche Ergebnisse sowie Zusammenarbeit mit anderen Forschungseinrichtungen**

Mit Projektende am 31.08.2025 konnte das im Projektverlauf entwickelte Verfahren zur Erfassung von Preisinformationen und weiteren relevanten Metadaten in den Produktivbetrieb überführt werden. Das Verfahren erlaubt eine qualitätsgesicherte Erfassung der Zeitschriften der drei Verlage ACS, Frontiers und Wiley. Parallel dazu wurde zum Projektende die Entwicklung des interaktiven Frontends abgeschlossen. Der ebenfalls im Projekt entwickelte Journal Price Index ermöglicht es, die Preisentwicklung der Publikationsgebühren nachzuzeichnen und greifbar zu machen. Über verschiedene Kommunikationskanäle (u.a. Mastodon) und Veranstaltungen wurde die Funktionsweise und Zielsetzung des Preismonitors beschrieben. Eine erste Fachpublikation wurde nach dem Projektende in der deutschsprachigen Zeitschrift o-bib eingereicht.

Das im Projekt entwickelte Verfahren zur KI-basierten Erfassung von Preis- und Metadaten aus Webseiten stößt bei vielen Akteuren im Handlungsfeld Elektronisches Publizieren auf ein großes Interesse und ermöglicht es, den im Projektnamen formulierten Anspruch eines »Publishing Analytics Tracker« zu konkretisieren. Über die Bereitstellung von reinen Preisdaten hinaus ist das Projekt PANTER ein wichtiger Beitrag für mehr Transparenz am akademischen Publikationsmarkt.

Bereits im Projektverlauf konnten erste Kooperationen mit anderen Projekten realisiert werden. Diese werden kontinuierlich ausgebaut und verstetigt. Hervorgehoben werden soll an dieser Stelle die Zusammenarbeit mit dem Projekt AdoReOA der UB Braunschweig, die sich im Nachfolgeprojekt PANDA manifestiert, in dessen Rahmen über ein gemeinsames Datenschema der Datenaustausch zwischen beiden Systemen vereinfacht und die Datenbasis von PANTER ausgebaut und erweitert wird.

Gefördert durch:



Bundesministerium  
für Forschung, Technologie  
und Raumfahrt

**HTW**D

Hochschule für Technik und  
Wirtschaft Dresden  
University of Applied Sciences



## PANTER – Publishing Analytics Tracker

<https://projekt-panter.de>

Sachbericht zum Verwendungsnachweis  
Teil II: Eingehende Darstellung (Stand 12.02.2026)

---

ZE:

Hochschule für Technik und Wirtschaft Dresden

Förderkennzeichen:

16KOA022B

---

Vorhabenbezeichnung:

**Verbundprojekt: Publishing Analytics Tracker – PANTER; Teilvorhaben B: Implementierung des Systems**

---

Laufzeit des Vorhabens:

**01.09.2023 - 31.08.2025**

---

Ansprechpartner:

**Prof. Dr.-Ing. Dietrich Kammer**

Professur Technische Visualistik, Fakultät Informatik/Mathematik  
Hochschule für Technik und Wirtschaft Dresden (HTWD)  
Hochschule für angewandte Wissenschaften  
kammer@htw-dresden.de

---

## Teil II: Eingehende Darstellung

Das Projekt wurde im Programm zur Etablierung einer gelebten Open-Access-Kultur in der Forschungs- und Wissenschaftspraxis des Bundesministeriums für Forschung, Technologie und Raumfahrt (BMFTR, vormals BMBF) für zwei Jahre gefördert. Es ist dem Themenfeld 3: »Förderung von Vorhaben, die die vielfältigen Bedarfe einer gelebten Open-Access-Kultur aufgreifen« zugeordnet. Das erklärte Förderziel des BMFTR war es, über das Programm weiterhin aktiv den Wandel wissenschaftlichen Publizierens hin zu mehr Open Access voranzutreiben und die Wissenschaft bei ihren diesbezüglichen Bestrebungen zu unterstützen. Als konkrete Zielsetzung für das Themenfeld 3 wurde folgende Zielsetzung benannt, die das Projektvorhaben adressiert: »Innovative Beiträge zur Etablierung einer gelebten Open-Access-Kultur in Deutschland überzeugend darstellen und dabei neue Akzente beim Wandel hin zu mehr Open Access setzen können«. Als Beispiele wurden in diesem Zusammenhang Vorhaben zur Entwicklung, Erprobung und Implementierung von Werkzeugen, Services oder Infrastrukturen genannt.

PANTER adressiert das Problem fehlender Preistransparenz auf dem Publikationsmarkt. Das Projekt zielt damit auf ein strukturelles Hemmnis sowohl für die Open-Access-Transformation als auch für die Innovationsfähigkeit des Publikationsmarktes. Insbesondere Wissenschaftlerinnen und Wissenschaftlern fehlt es bislang an übersichtlichen Informationen zu den finanziellen Bedingungen des Publizierens. Dies schränkt die wissenschaftliche Selbstverantwortung und Wahlfreiheit ein, stützt monopolartige Strukturen und wirkt sich in Folge bremsend auf die Weiterentwicklung von Dienstleistungen und Produkten aus. Der Preismonitor als öffentlich zur Verfügung stehendes Online-Tool erlaubt es den unterschiedlichen Stakeholdern des wissenschaftlichen Publizierens, Preise und Marktentwicklungen zu verfolgen. Mit seiner Hilfe können Forschende, Wissenschaftseinrichtungen und Bibliotheken ihre Entscheidungen und Aktivitäten finanzierungsbewusst belegen und somit die Kosteneffizienz der Finanzierung von Open-Access-Publikationen deutlich erhöhen. PANTER leistet mit dem Werkzeug Preismonitor daher einen wichtigen Beitrag zu einem nachhaltigen Open-Access-Ökosystem.

Vom Projekt PANTER profitieren verschiedene Zielgruppen:

- Forschende (Vergleich verschiedener Zeitschriften und Preissensibilisierung)
- Forschung (Valide Datenbasis für Untersuchungen des Publikationsmarktes und der Preispolitik von Verlagen)
- Hochschulverwaltung und -bibliotheken (Prognosen Mittelbedarf, Verhandlungsgrundlage für Verlagsverträge)
- Wissenschaftspolitik (Valide Datenbasis für wissenschaftspolitische Entscheidungen)

Im Betrieb soll künftig der Beitrag des Projektes zur Zielsetzung des BMFTR-Programmes entlang folgender quantitativer Indikatoren bemessen werden:

- Nutzungszahlen/Zugriffe pro Monat
- Impact in Wissenschaftskommunikation (Rezeption in Social Media und auf Fachbeiträge)

## II.1 Wichtigste Positionen des zahlenmäßigen Nachweises

Durch die Mittel in Position 0812 (Beschäftigte E12-E15) konnte von April 2024 bis August 2025 ein wissenschaftlicher Mitarbeiter an der HTW Dresden mit einer halben Stelle beschäftigt werden, um die Konzeptions- und Umsetzungsarbeiten für den Preismonitor wesentlich durchzuführen und zu begleiten. Die verzögerte Einstellung (geplant war Januar 2024) zum Projektbeginn wurde durch den ständigen Austausch des Projektleiters an der HTW Dresden mit den anderen Projektpartnern sowie einer Komprimierung der Arbeitspakete an der HTW Dresden kompensiert.

Der wissenschaftliche Mitarbeiter wurde durch eine wissenschaftliche Hilfskraft im Zeitraum von September 2024 bis August 2025 unterstützt, die durch die Mittel in Position 0822 (Beschäftigungsentgelte) finanziert wurde. Weitere wesentliche Arbeiten an der Pipeline für die Datenerfassung sowie am Frontend und Backend des Preismonitors wurden durch die Mittel in Position 0835 durch die Vergabe von vier Aufträgen an verschiedene Unterauftragnehmer realisiert. Durch die Mittel in Position 0846 konnten die notwendigen Dienstreisen zur Bekanntmachung und wissenschaftlichen Veröffentlichung auf Konferenzen und Veranstaltungen sichergestellt werden.

## II.2 Notwendigkeit und Angemessenheit der geleisteten Projektarbeiten

Der Entwicklung des Preismonitors wurden einige grundsätzliche Anforderungen vorangestellt:

- Einfache und verlässliche Nutzbarkeit
- Verschiedene Interaktionsmöglichkeiten für Nutzende
- Vergleiche zwischen Preisen und deren Entwicklung auf verschiedenen Aggregierungsebenen
- Entwicklung eines Preisindex (Abbildung der Preisentwicklung)
- Robuste und skalierbare Technologien zur Preisdatenerfassung
- Hohe Datenqualität/Datenintegrität, die durch Qualitätstests validiert werden
  - *Vollständigkeit*: Vollständige Erfassung der Preisdaten inklusive der individuellen Rabattoptionen und Zuschläge aller Zeitschriften eines Verlages
  - *Aktualität*: Regelmäßige Updates, um den Datenbestand aktuell zu halten
  - *Genauigkeit*: Korrektheitsgrad von über 90 Prozent
  - *Konsistenz*: Gleiche Informationen werden in demselben Format abgebildet
  - *Einzigartigkeit*: Vermeidung von redundanten Datensätzen (Duplikaten)
  - *Gültigkeit*: Zeitlich determinierte Informationen werden in ihrem Gültigkeitsbereich mit Zeitstempeln ausgewiesen

Das methodische Vorgehen des Vorhabens orientierte sich stark am Vorgehensmodell des Benutzerzentrierten Designs (User-Centered Design, UCD). Dieser Prozess konzentriert sich auf die Bedürfnisse, Verhaltensweisen und User Experience (UX) der Nutzenden des Preismonitors. Dabei wurden mehrere Phasen durchlaufen:

- **Define**: In dieser Phase bestimmte das Projektteam die Ziele des Projekts sowie die Zielgruppe und die wichtigsten Aufgaben, die der Preismonitor unterstützen soll.
- **Research**: In dieser Phase sammelte das Team Informationen über die Zielgruppe sowie den Kontext, in dem der Preismonitor verwendet werden soll. Von zentraler Bedeutung für das Vorhaben war der Zugang zu historischen und aktuellen Preisinformationen. Diese werden von diversen Akteuren in unterschiedlicher Auflösung und Qualität bereitgestellt. Die

Anbieter stellen ihre Preisinformationen entweder auf ihren Webseiten oder in Form von Listen bereit, die in dieser Phase ermittelt wurden.

- **Ideate:** In diesem Prozessschritt wurde in einem ersten Schritt das relationale Datenmodell zur Erfassung der relevanten Daten entworfen. Im weiteren Verlauf dieser Phase entwickelte das Team Ideen für das Design des Preismonitors.
- **Prototype:** In dieser Phase erstellte das Team einen Prototyp des Produkts, d. h. eine vereinfachte Version, die mit Nutzenden getestet werden konnte. Zudem konnten hier vom Projektteam die ersten Ergebnisse aus dem Workflow Datenerfassung getestet werden.
- **Test:** In dieser Phase führte das Team weitere Tests durch, um Feedback zum Prototyp zu erhalten. Dabei wurden insbesondere die Benutzererfahrung (User Experience, UX), die Benutzerfreundlichkeit (Usability), die Wirksamkeit des Designs (emotionale Reaktionen auf die Ästhetik des Monitors) erfasst.
- **Iterate:** Auf Grundlage des Feedbacks aus den Tests iterierte das Team den Entwurf und nahm Änderungen und Verbesserungen vor. Ebenfalls in dieser Phase erfolgte die weitere Validierung und Evolution der Datenbank.
- **Implement:** In dieser letzten Phase setzte das Team den endgültigen Entwurf um und bereitete den Preismonitor für die Einführung vor. Dies umfasste die Bereitstellung der Produktivdatenbank, die Backend- und Frontend-Programmierung sowie die Definition von Tests und der Qualitätssicherung.

Im Ergebnis dieses Prozesses wurde ein datenbankgestützter Webservice entwickelt, dessen Architektur auf den in Abbildung 1 visualisierten Komponenten basiert.

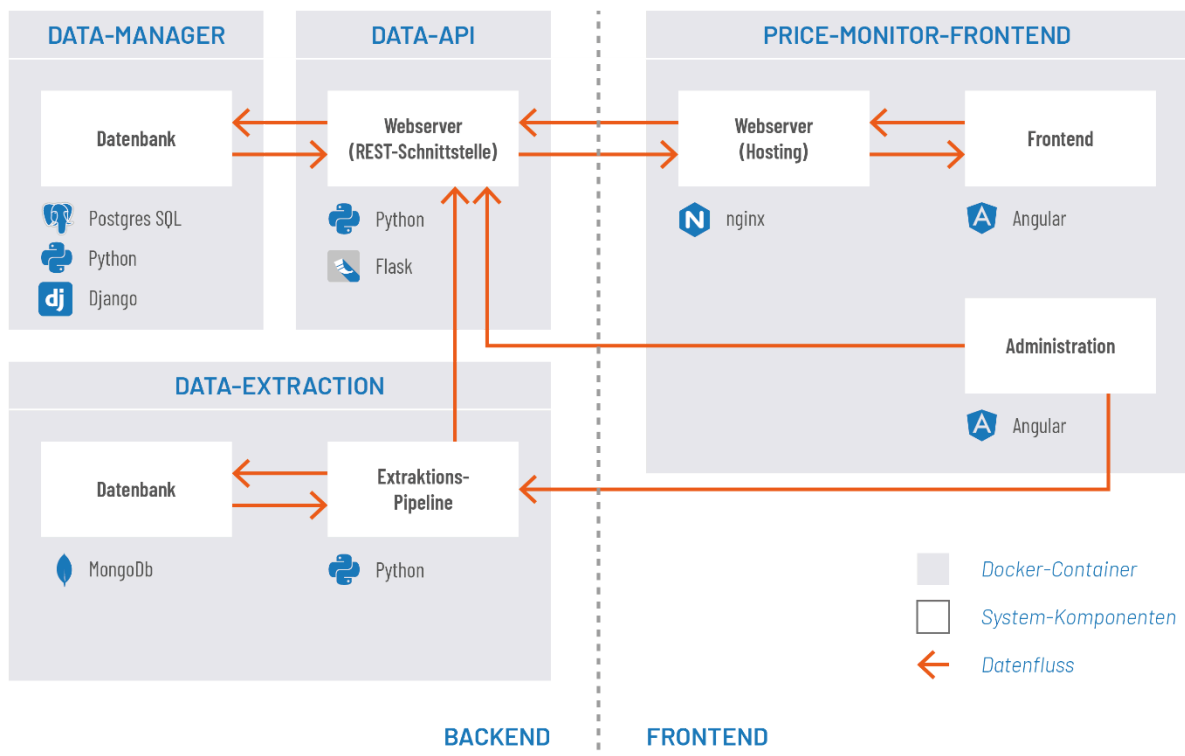


Abbildung 1: Tech-Stack von PANTER

Der Tech-Stack besteht aus vier zentralen Komponenten:

1. *Data-Extraction*: Eine Extraktionspipeline mit zugehöriger dokumentenorientierter MongoDB-Datenbank, in der die aus den Webseiten extrahierten Rohdaten erfasst, strukturiert und aufbereitet werden.
2. *Data-Manager*: Eine relationale PostgreSQL-Datenbank, die die aufbereiteten Daten im in AP3 erarbeiteten Datenschema speichert..
3. *Price-Monitor-Frontend*: Ein Webserver, der sowohl das öffentlich zugänglicher Frontend des Preismonitors als auch ein zugriffsbeschränktes Administrations-Webinterface hostet und
4. *Data-API*: Ein Webserver mit der REST-API, die die einzelnen Komponenten verbindet.

Vereinfacht ausgedrückt funktioniert der Workflow wie folgt:

1. Eine zuvor definierte Menge von Webseiten wird scriptbasiert ausgelesen (Webscraping) und der HTML-Quelltext dieser Seiten anschließend in der Mongo-DB abgespeichert.
2. Mittels eines großen Sprachmodells (LLM) werden aus dem HTML-Quelltext alle relevanten Daten und Informationen extrahiert, unifiziert und in eine strukturierte Form gebracht.
3. Über die REST-API werden diese aufbereiteten Daten dann in die PostgreSQL-Datenbank überführt.
4. Das Webfrontend greift ebenfalls über die REST-API auf die Daten in der Datenbank zurück und erlaubt den Nutzern vielfältige Interaktionsmöglichkeiten zur Ansicht der relevanten Daten.
5. Die einzelnen Prozesse können über eine intern für das Projektteam zugängliche Administrationsoberfläche gesteuert werden (ebenfalls primär über die REST-API).

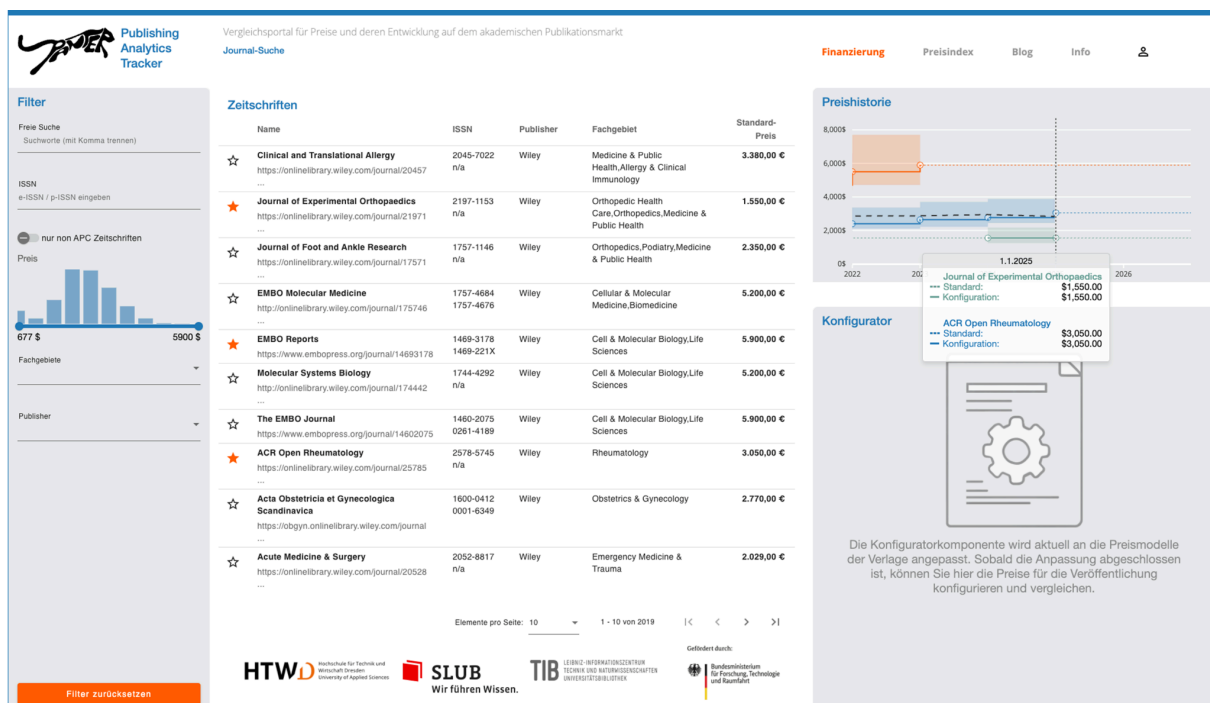


Abbildung 2: PANTER-Frontend

Der Zielgruppe von PANTER bleiben die Prozesse im Backend verborgen. Für sie ist das Frontend mit seinen vielfältigen Interaktionsmöglichkeiten von Bedeutung. Grundsätzlich ist das Frontend in drei

Fensterbereiche unterteilt, die sich von links nach rechts anordnen. Der rechte Fensterbereich ist noch einmal in einen oberen und unteren Bereich unterteilt (Abbildung 2).

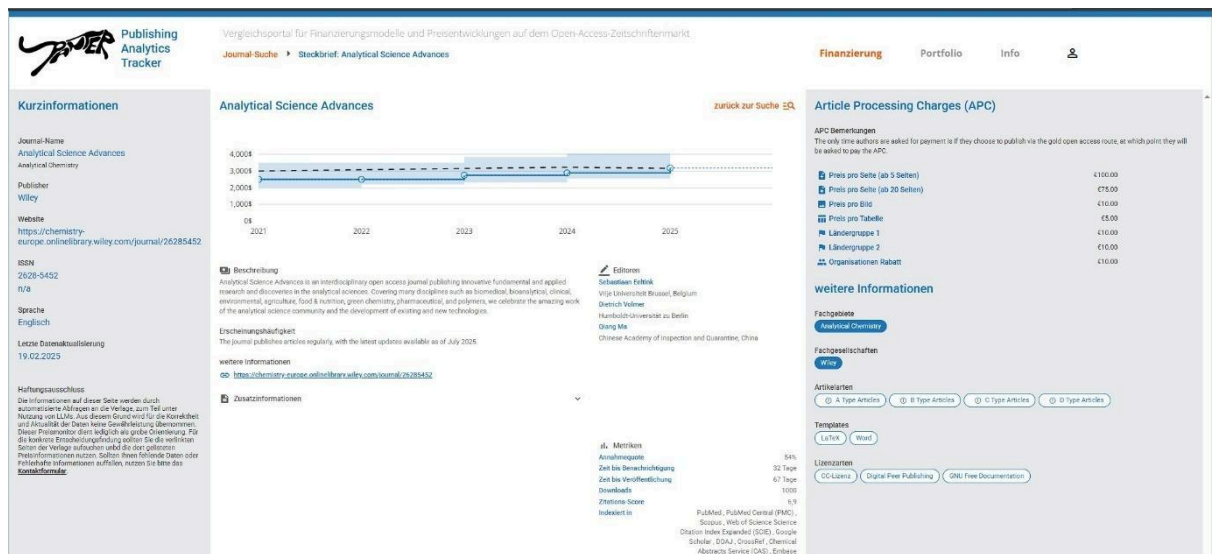


Abbildung 3: Journalsteckbrief

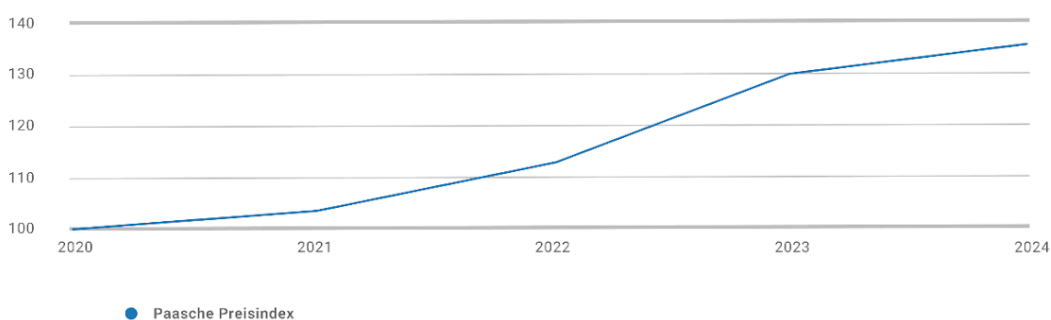
Der linke Fensterbereich bietet vielfältige Filtermöglichkeiten. Über Freitextfelder, ISSN, Preisspannen, Fachgebiete oder Verlage können in einem ersten Selektionsschritt Journale recherchiert werden. Diese Auswahl an Journalen wird anschließend mit einigen relevanten Kenngrößen im mittleren Fenster angezeigt. Wenn die Nutzenden eine oder mehrere Journale über das Sternensymbol markieren, erscheinen im rechten geteilten Fenster weitere Informationen bzw. Interaktionsmöglichkeiten. Im oberen Bereich des rechten Fensters wird die Preisentwicklung der Publikationsgebühren über die letzten Jahre visualisiert. Wenn mehrere Journale ausgewählt wurden, wird diese Preisentwicklung vergleichend dargestellt. Im unteren rechten Fenster wird es in Zukunft ein sogenannter Konfigurator ermöglichen, die Preise zu individualisieren, indem die Nutzenden zusätzliche preisbeeinflussende Informationen wie die Seitenzahl oder Mitgliedschaften in Fachgesellschaften setzen können. Ähnlich wie im Projekt B!SON gibt es auch einen sogenannten Journalsteckbrief (siehe Abbildung 3), der aber nicht wie in B!SON die Daten aus dem DOAJ abbildet, sondern aktuelle Daten ausweist, die über die Datenextraktion aus den Webseiten gewonnen wurden. Ein Teil dieser Daten, wie etwa die Dauer des Review-Prozesses, die Namen der Editoren oder die Indexierung in bibliographischen Datenbanken wird in keinem anderen frei zugänglichen Nachweissystem ausgewiesen. Insbesondere für OA-Professionals interessant ist der Journal Price Index, der die Entwicklung der Publikationsgebühren als Index abbildet (Abbildung 4).

Die durchgeführten Arbeiten waren für die Erfüllung der Aufgaben notwendig und angemessen, da ein vergleichbarer Dienst unter gleichen Bedingungen mit den hier entwickelten Technologien nicht zur Verfügung steht. Die durch die HTW Dresden durchgeführten Arbeiten hätten ohne die bereitgestellte Förderung nicht durchgeführt werden können. Die erzielten Ergebnisse des Vorhabens und die bei der Projektbearbeitung gewonnenen Erkenntnisse zu KI-gestützten Verfahren des Webscrapings bieten der wissenschaftlichen Community über den unmittelbaren Anwendungszweck hinaus aufgrund ihrer Praxisrelevanz einen unmittelbaren Mehrwert und hohes Nutzungspotential. Die Rückmeldungen während der Projektlaufzeit waren sehr positiv.

Auch nach Projektende steht der wissenschaftlichen und bibliothekarischen Community ein erfolgreich getesteter Preismonitor zur Nutzung bereit, der über Folgeprojekte ausgebaut werden

soll. Sämtliche Projektarbeiten dienten dem bedarfsadäquaten Aufbau des Dienstes. Die Bereitstellung der Software, des Quellcodes und der Schnittstellen ermöglicht eine Nutzung durch weitere Einrichtungen. Der Quellcode für die Pipeline zur Datenerfassung soll nur ausgewählten Einrichtungen bereitgestellt werden. Eine Veröffentlichung unter einer freien Lizenz ist für diesen Teil des Quellcodes nicht geplant, um die Funktionalität des Verfahrens nicht zu gefährden. Im Projektverlauf erschwerten immer neue Schutzmechanismen das Webscraping. Eine Offenlegung des Quellcodes würde es ermöglichen, weitere Schutzmechanismen zu installieren und somit die Datenerfassung zu verhindern. Über ein bereits bewilligtes Folgeprojekt unter dem Namen PANDA mit dem neuen Projektpartner TU Braunschweig; diverse Präsentationen des Projektes sowie eingereichte bzw. geplante Publikationen ist ein umfangreicher Transfer der Projektergebnisse zu erwarten.

### Journal Price Index



Indikator	2020	2021	2022	2023	2024
<b>Paasche Preisindex</b>	100	103	112	130	136
Vergleich zum Vorjahr		+ 3%	+ 8,74%	+ 16,07 %	+ 4,62%

Abbildung 4: Journal Price Index

Im Folgenden werden die an der HTW Dresden im Teilvorhaben durchgeführten Arbeiten anhand der entsprechenden Arbeitspakete erläutert.

### AP 1 Anforderungsanalyse

Die HTW Dresden beteiligte sich beim Auftaktworkshop in Dresden am 22.02. und 23.02.2024 mit allen Projektbeteiligten zur Festlegung der Anforderungen und des weiteren Vorgehens. Der Beitrag der HTW Dresden war insbesondere die Moderation des User-Centered-Design-Prozesses und die Diskussion zum Design des Preismonitors. Als wesentliches Ergebnis entstanden für die weitere Arbeit die entsprechenden Entwurfsdokumente einschließlich Personas und User-Story-Maps, die in einem geteilten Miro-Board dokumentiert und fortlaufend weiterentwickelt wurden. In einem weiteren Projektworkshop in Dresden am 26. und 27.08.2024 wurden insbesondere die Designentwürfe und die aktuelle Umsetzung des AP 6 (Webservice und Interface) präsentiert und diskutiert.

### AP 3 Datenbankimplementierung und Betrieb

Im Rahmen dieses APs (AP 3.1) implementierte die HTW Dresden das Datenbankmodell aus AP 2 (siehe Sachbericht der SLUB Dresden) in einem offenen relationalen Datenbankmanagementsystem (DBMS) mit Hilfe der PostgreSQL-Software und überführte diese auf einer virtuellen Maschine im

VM-Cluster der HTW Dresden in den Betrieb (AP 3.2). Die Grunddaten aus AP 4 (siehe Sachbericht der SLUB Dresden) konnten übertragen sowie die in AP 5 automatisch erfassten gespeichert werden. Gemäß des User-Centered-Designprozesses wurde die Datenbank iterativ validiert und bis zum Ende des Projekts evolutioniert. Damit konnte der Meilenstein M 3.1 mit einem lauffähigen Datenbankserver erfolgreich abgeschlossen werden.

### AP 5 Workflow Datenerfassung

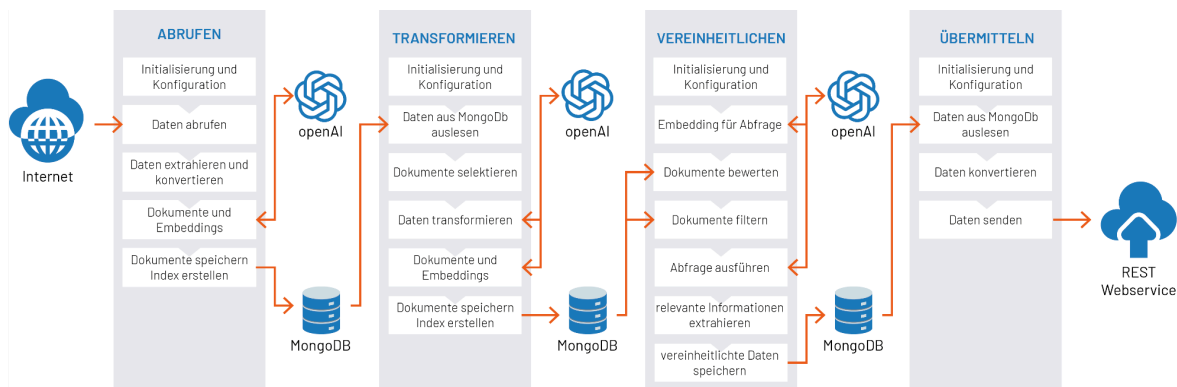


Abbildung 5: Datenerfassungsworkflow mit Hilfe der verwendeten Webscraping-Pipeline basierend auf Python und dem Einsatz von Large-Language-Models.

In diesem Arbeitspaket wurde der Meilenstein 5.1 zur Dokumenten-Spezifikation abgeschlossen sowie der Datenimport-Workflow mit Hilfe einer bereits verfügbaren Webscraping-Pipeline basierend auf Python und Large-Language-Models umgesetzt (siehe Abbildung 5). Dieser Ansatz wurde exemplarisch für eine Auswahl an Verlagen erprobt. Die Datenerfassung wurde in einem eigenen Code-Repository umgesetzt und ließ sich durch die Virtualisierungssoftware Docker auf dem Zielsystem an der HTW Dresden entsprechend aufsetzen. Damit konnte im Projekt ein innovativer und zukunftsfähiger Ansatz im Bereich des Webscrapings eingebunden werden, der nach ausführlicher Prüfung und Validierung modifiziert und erweitert wurde. Dafür wurden die zunächst gesparten Personalmittel für den im Zusammenhang mit diesem Arbeitspaket vorgesehenen Unterauftrag verwendet.

### AP 6 Web Service & Interface

Zunächst wurde die erste Iteration für das PANTER-System durchgeführt. Sowohl die Konzeption (AP 6.1), das Backend (AP 6.2) als auch das Frontend (AP 6.3) wurden dabei umgesetzt und auf einer virtuellen Maschine als Testversion allen Projektpartnern zur Verfügung gestellt. Die Projektergebnisse der Konzeption (Moodboards, Wireframes, Technologie-Stack) wurden über Miro, Figma und Github dokumentiert, und zur Unterstützung der Umsetzung wurden Programmierarbeiten an externe Dienstleister vergeben. Damit stand für erste Benutzungstests fristgerecht eine Alpha-Version (M 6.1) bereit, die bereits eine passfähige Autorisierungslösung für das Backend sowie eine erste Interaktionsoberfläche zur Datensichtung bot.

In der darauf folgenden Entwicklungsphase hin zur Beta-Version (M 6.2) wurde der Funktionsumfang gemäß den Antragszielen signifikant erweitert. Dies umfasste die Implementierung der spezifischen Visualisierungsmetaphern, Listenansichten sowie komplexer Filtermechanismen (z. B. nach Journalname, ISSN oder Verlag) zur Exploration und zum Export des Datenbestands. Zur Evaluierung der Nutzung wurde zudem das geplante Datennutzungs- und Logging-Konzept integriert und getestet.

Den Abschluss der Arbeiten bildete das finale Release (M 6.3), bei dem der Web-Service inklusive der Datenbank und der integrierten Programmierschnittstelle (aus AP 7) erfolgreich in die Infrastruktur der HTW Dresden überführt wurde. Der Monitor wird dort nun zentral gehostet und betrieben, womit die technische Basis geschaffen wurde, um das Interface für die geplante Dauer von zunächst fünf Jahren aufrechtzuerhalten.

#### **AP 7 Programmierschnittstelle**

Das Arbeitspaket 7 diente der Schaffung einer Programmierschnittstelle, um die Projektdatenbank (AP 3) auch unabhängig vom grafischen Interface zugänglich zu machen. Die Schnittstelle wurde zunächst als REST-Architektur definiert, womit der Meilenstein M 7.1 (API-Definition) bereits in der frühen Projektphase erreicht wurde. Die anschließende technische Umsetzung (AP 7.2) wurde durch die Vergabe von Unteraufträgen an externe Dienstleister unterstützt. Die entwickelte Lösung fungiert dabei als Middleware zwischen der Datenerfassungspipeline und dem Datenbankserver und wird in einem gesonderten Code-Repository verwaltet.

Im weiteren Verlauf lag der Fokus auf der Absicherung und Nutzbarmachung der API. Hierfür wurde ein Sicherheitskonzept auf Basis von JSON Web Tokens (JWT) implementiert und eine umfassende Online-Dokumentation erstellt, die Eingabefelder und Ausgabeformate transparent darlegt. Mit der erfolgreichen Integration in die Hosting-Umgebung der HTW Dresden konnte schließlich das API-Deployment (M 7.2) abgeschlossen werden, sodass externen Akteuren nun eine stabile und dokumentierte Schnittstelle für automatisierte Anfragen zur Verfügung steht. Die Freigabe der Nutzung unterliegt einer entsprechenden Governance-Struktur, die nach Projektende geschaffen wird (siehe oben und Abschnitt II.3).

### **II.3 Voraussichtlicher Nutzen**

Der im Projektverlauf erarbeitete Preismonitor steht als Prototyp mit Preis- und weiteren relevanten Metadaten von insgesamt drei großen Wissenschaftsverlagen offen und kostenfrei zur Verfügung und trägt dazu bei, eine der großen Informationslücken, die einer nachhaltigen und innovationsfreudigen Open-Access-Transformation im Wege stehen, zu schließen. Dieser Lückenschluss ist auch deshalb von großer Bedeutung, da der von der Coalition S betriebene Journal Comparison Service Ende April 2025 eingestellt wurde (siehe Abschnitt II.4).

Die Projektpartner SLUB und TIB bekennen sich dazu, alle nötigen und für den Betrieb des modularen Systems nötigen Leistungen mindestens für fünf Jahre nach Projektende in Eigenleistung zu erbringen und die Ergebnisse interessierten Partneereinrichtungen und OA-Initiativen wie dem DOAJ bereitzustellen. Dieses Bekenntnis deckt sowohl die aktive Datenakquise – und hier vor allem die Redaktion der Daten aus dem Webscraping – als auch die Bereitstellung des Journal Price Index ab. Die HTWD erklärt sich bereit, den Preismonitor bis zu fünf Jahren nach Projektende zu aktualisieren und die Schnittstellen zu pflegen. Es ist die Absicht der Projektpartner, diesen Beitrag zu einer offenen Publikationslandschaft langfristig bereitzustellen.

In den an das Projektende anschließenden fünf Jahren wird neben dem Aufbau einer Governance-Struktur für den Produktivbetrieb und die Weiterentwicklung der Scraping-Pipeline in einem Netzwerk auch ein Finanzierungsmodell erarbeitet, um zumindest einen Teil der anfallenden Betriebskosten abzudecken. Durch den unter einer offenen Lizenz (AGPL bzw. GPLv3) verfügbaren Quellcode kann das System auch durch Dritte angeboten und weiterentwickelt werden. Die im Projekt gewonnenen Erkenntnisse und entwickelten Technologien sollen in weitere Vorhaben

münden und bieten das Potential für neue Dienste und weitere Drittmittelvorhaben. Entsprechende Kooperationen mit anderen Projekten wurden entweder bereits in Form eines bereits bewilligten neuen Projektes realisiert (PANDA) oder sind mit nationalen bzw. internationalen Akteuren angedacht.

Eine Kooperation mit der Forschungsgruppe um Stefanie Haustein (University of Ottawa) erscheint besonders sinnvoll, da diese Projektgruppe über viele Jahre Erfahrung im Preismonitoring und der Analyse dieser Daten verfügt und über entsprechende Publikationen eine hohe Sichtbarkeit in diesem Handlungsfeld erarbeitet hat (Haustein, 2024; Butler, 2024). Über die Datenerfassungspipeline des Projektes PANTER konnten zusätzliche Preisdaten erhoben werden, die der University of Ottawa bereitgestellt wurden. Neben dieser leistungsstarken Pipeline sind auch das interaktive Frontend sowie der Journal Price Index Alleinstellungsmerkmale des Projektes PANTER. Wie bereits erwähnt, wurde der Journal Comparison Service der Coalition S im Frühjahr 2025 eingestellt. Weitere vergleichbare Dienste sind nicht bekannt.

#### **II.4 Während der Durchführung des Vorhabens dem Zuwendungsempfänger bekannt gewordenen Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen**

Der im Projektantrag genannte »Journal Comparison Service« (JCS) der internationalen Open Access Initiative »cOAlition S« stellt zum April 2025 den Betrieb ein. Dies hat zwar keine direkten Auswirkungen auf PANTER. Es unterstreicht aber die großen Herausforderungen, die bei der Herstellung von Preistransparenz zu bewältigen sind, sowie die Bedeutung des verlagsunabhängigen Ansatzes, den PANTER verfolgt. Der JCS existierte seit 2021. Er stellte Preis- und Kosteninformationen für eine eng gefasste Nutzergruppe nach aufwendigem Registrierungsprozess bereit und war dazu auf freiwillige Datenlieferungen der Verlage angewiesen. cOAlition S begründete die Einstellung des Service mit einem negativen Kosten-Nutzen-Verhältnis und schlechter Beteiligung der Verlagsseite. Weitere vergleichbare Dienste sind nicht bekannt.

#### **II.5 Erfolgte oder geplante Veröffentlichungen des Ergebnisses**

Während der Projektlaufzeit wurde PANTER mehrfach auf nationalen Konferenzen präsentiert. Auch über die Projektförderung hinaus soll der Preismonitor weiterhin der interessierten Öffentlichkeit vorgestellt werden. Fachspezifische Publikationen in deutscher und englischer Sprache werden die Bekanntheit des Preismonitors flankierend weiter steigern.

Die technischen Arbeiten der HTW Dresden wurden auf einem Workshop mit zugehörigen Proceedings vorgestellt:

- Pechstein, G., Müller, M., & Kammer, D. (2025). Web based pipeline administration and Docker remote control for extracting open access journal metadata. Proceedings of the 28th Bilateral Student Workshop HTW Dresden and CTU Prague – User Interfaces & Visualization *(to appear on qucosa)*

Die entstandene Software des Preismonitors steht durch das Hosting in der technischen Infrastruktur der HTW Dresden nachhaltig zur Verfügung:

- Webfrontend: <https://projekt-panter.de/>

- API: <https://projekt-panter.de/api/pricemonitor/>

Aus Sicherheitsgründen ist die Swagger-Dokumentation der API auf dem Server deaktiviert.

Eine Beschreibung der API-Routen kann über die folgende URL abgerufen werden:

<https://editor.swagger.io/?url=https://projekt-panter.de/api/pricemonitor/swagger.json>

- Quellcode unter AGPL-Lizenz: <https://github.com/panter-project/> (nur ausgewählte Teile der Umsetzungen sind öffentlich einsehbar, siehe oben)