



# Planung von APCs – Kann KI uns weiterhelfen?

Open-Access-Tage 2025

18.09.2025

Constantin Lehenmeier, Dr. Gernot Deinzer  
Universitätsbibliothek Regensburg



Universität Regensburg



Vorüberlegungen

# **INFORMATIONSBUDGET**

# Informationsbudget

„Unter dem Begriff Informationsbudget wird hier ein Instrument des Finanzmanagements verstanden, mit dem alle Einnahmen und Ausgaben für die wissenschaftlichen Informationen bewirtschaftet werden.“

Pampel, H. (2019): Auf dem Weg zum Informationsbudget: zur Notwendigkeit von Monitoringverfahren für wissenschaftliche Publikationen und deren Kosten; Arbeitspapier, Potsdam : Helmholtz Open Science Koordinationsbüro, 15 p.  
<https://doi.org/10.2312/os.helmholtz.006>

# Umsetzung Universität Regensburg

## Zentrale Bearbeitung und Erfassung Kauf-; Lizenz- Und Publikationskosten:

Deinzer, Gernot (2023) Implementierung des Informationsbudgets an der Universität Regensburg: Umsetzung, Strukturen und Geschäftsprozesse, Open Access Tage 2023, Berlin, Folien: DOI [10.5283/epub.54763](https://doi.org/10.5283/epub.54763), Video: DOI: [10.5446/66725](https://doi.org/10.5446/66725)

Deinzer, Gernot und Lang, Cornelia (2024) All in One: Informationsbudget an der Universität Regensburg: Umsetzung, Strukturen und Geschäftsprozesse. Zeitschrift für Bibliothekswesen und Bibliographie ZfBB 71 (2), S. 91-101. DOI: [10.5283/epub.57999](https://doi.org/10.5283/epub.57999) (CC-BY-Lizenz)

# **WIE KÖNNEN WIR ZUKÜNFTIGE KOSTEN ABSCHÄTZEN?**



Universität Regensburg

**Dr. Gernot Deinzer**  
Open Access Beauftragter  
Universitätsbibliothek Regensburg

Rückblick

# KOSTENGRENZEN

# Abschätzung mit Grenzen:

## **Obere Grenze**

**Artikel, für welche Publikationsgebühren fällig sind**

Unabhängig von dem corresponding author

## **Schätzung**

**Artikel bezahlt von UR**

Durchschnittspreis

## **Untere Grenze:**

**Artikel, welche von UR bezahlt werden müssen**

Gold Open-Access-Artikel, bei denen alle Autor:innen die Affiliation UR haben

## Grenzen bezüglich Affiliation

$$\begin{aligned} \sum_{Pubp} w_{p,onlyUR} * (R_p * PG_p + \Theta_p * SP) &< \\ \sum_{Pubp} w_p * (R_p * PG_p + \Theta_p * SP) & \\ < \sum_{Pubp} w_{p,all} * (R_p * PG_p + \Theta_p * SP) \end{aligned}$$

$w_{p,onlyUR}$  : Gewicht (=1 nur UR Autor\*innen, 0 sonst)

$w_{p,all}$  : Gewicht (=1 kostenpflichtig, 0 sonst)

$R_p$  : Rabatt für Publikation  $p$

$PG_p$  : Publikationsgebühr Publikation  $p$

$SP$  : Servicepauschale

$\Theta_p$  : 1 für Publikation mit Servicepauschale, 0 sonst



Ziel

# **VORHERSAGE APC UND PUBLIKATIONSVERHALTEN**

# Herausforderung

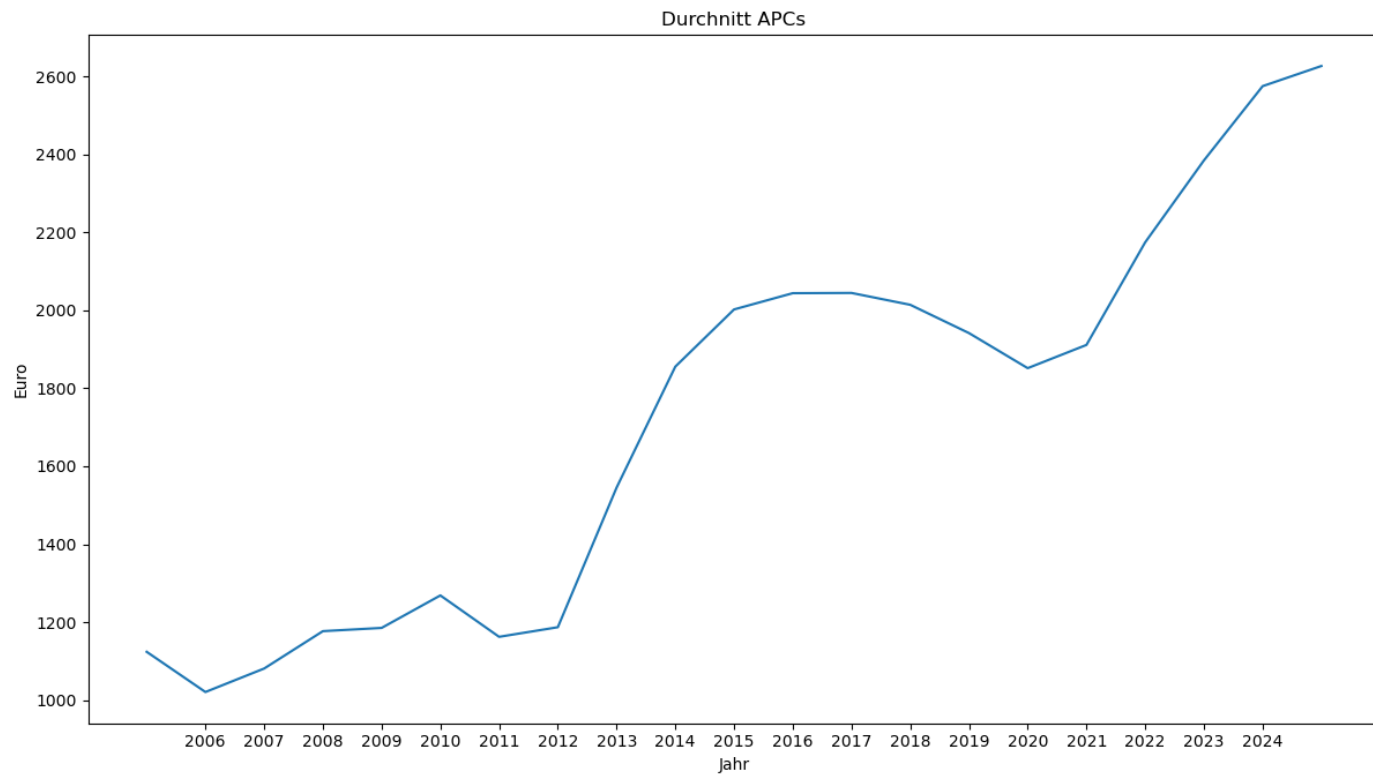
$$\sum_{Pubp} w_p * (R_p * PG_p + \Theta_p * SP)$$

Wie kann man die

- Anzahl der Veröffentlichungen
- Gebühren dieser Veröffentlichung
- Und eventuelle Rabatte

Effektiv abschätzen?

# Entwicklung APCs



# **Maschinelles Lernen**

## **Erfolge:**

- **Vorhersage Aktienkurs**
- **Vorhersage Wetter**
- **Prognose Kaufverhalten von Kunden**
- **Immobilienpreise**

Frage

**KANN MAN MASCHINELLES  
LERNEN ZUR VORHERSAGE  
VON APCS EINSETZEN?**



Universität Regensburg

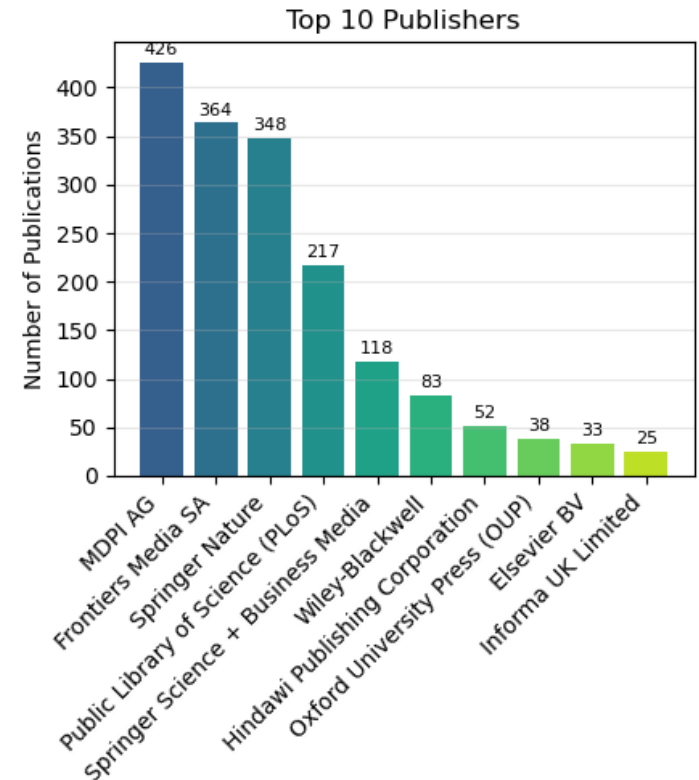
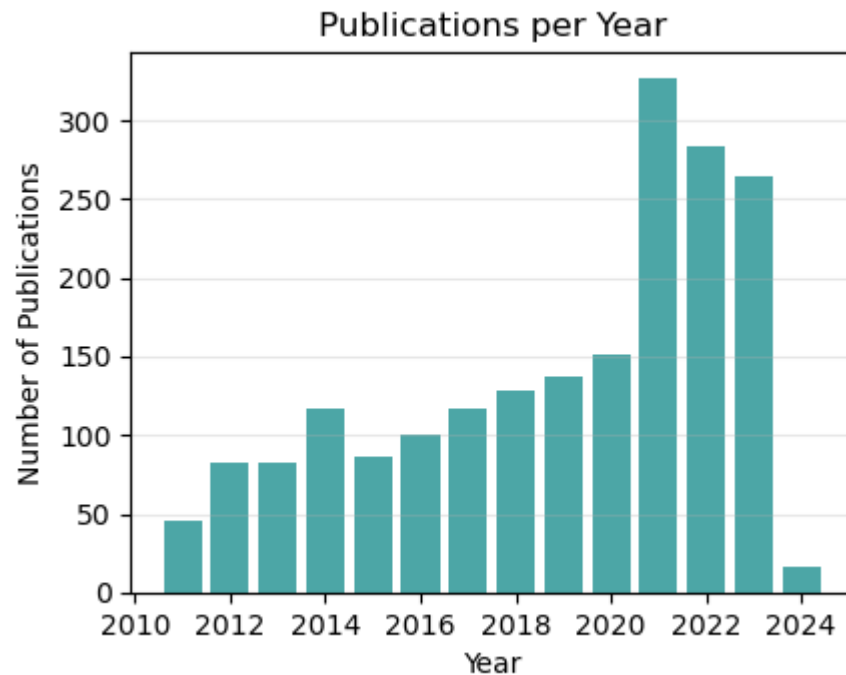
**Dr. Gernot Deinzer**  
Open Access Beauftragter  
Universitätsbibliothek Regensburg

Uni Regensburg

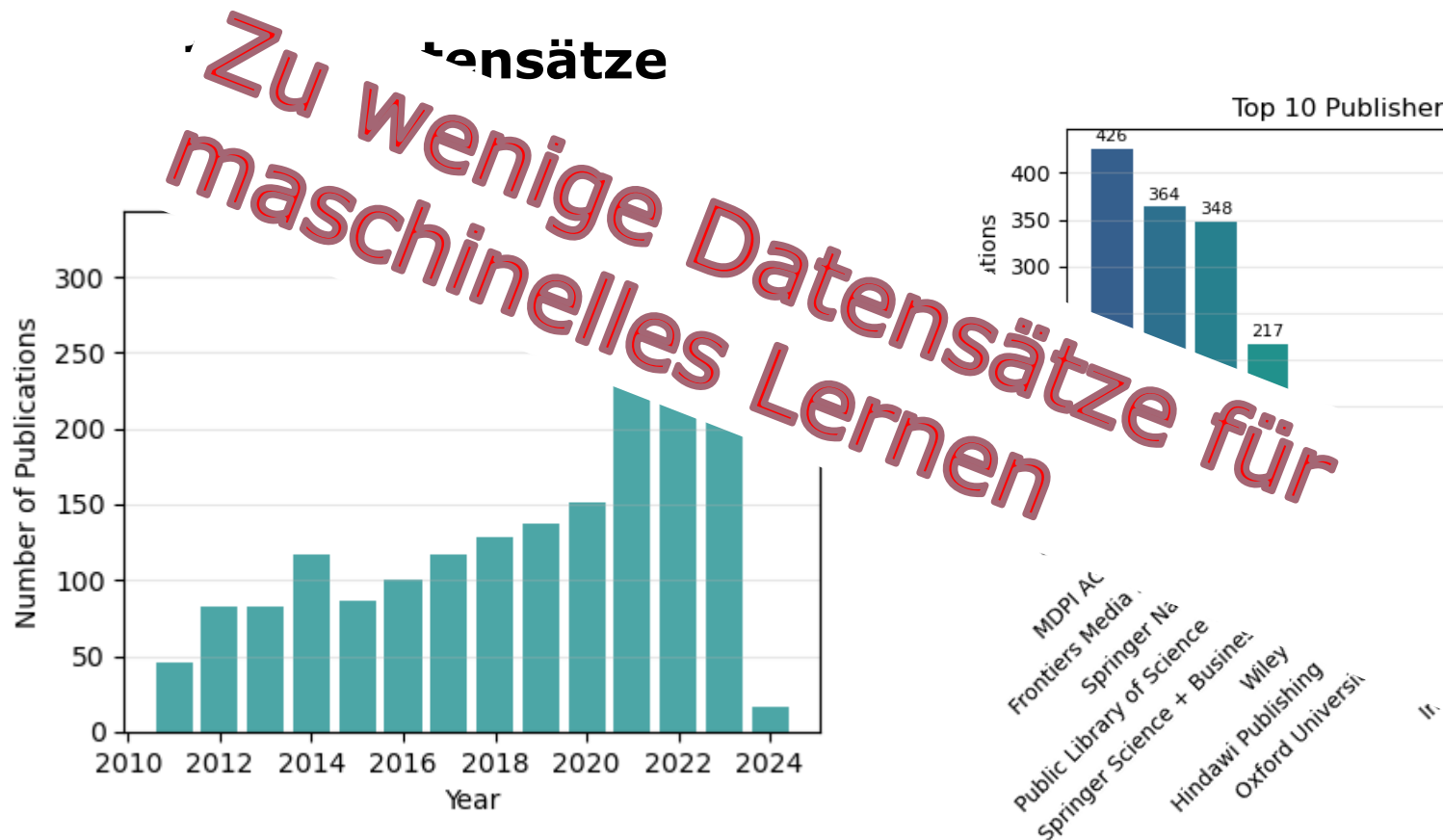
**BEZAHLTE GEBÜHREN**

# Übersicht Universität Regensburg

## 1941 Datensätze

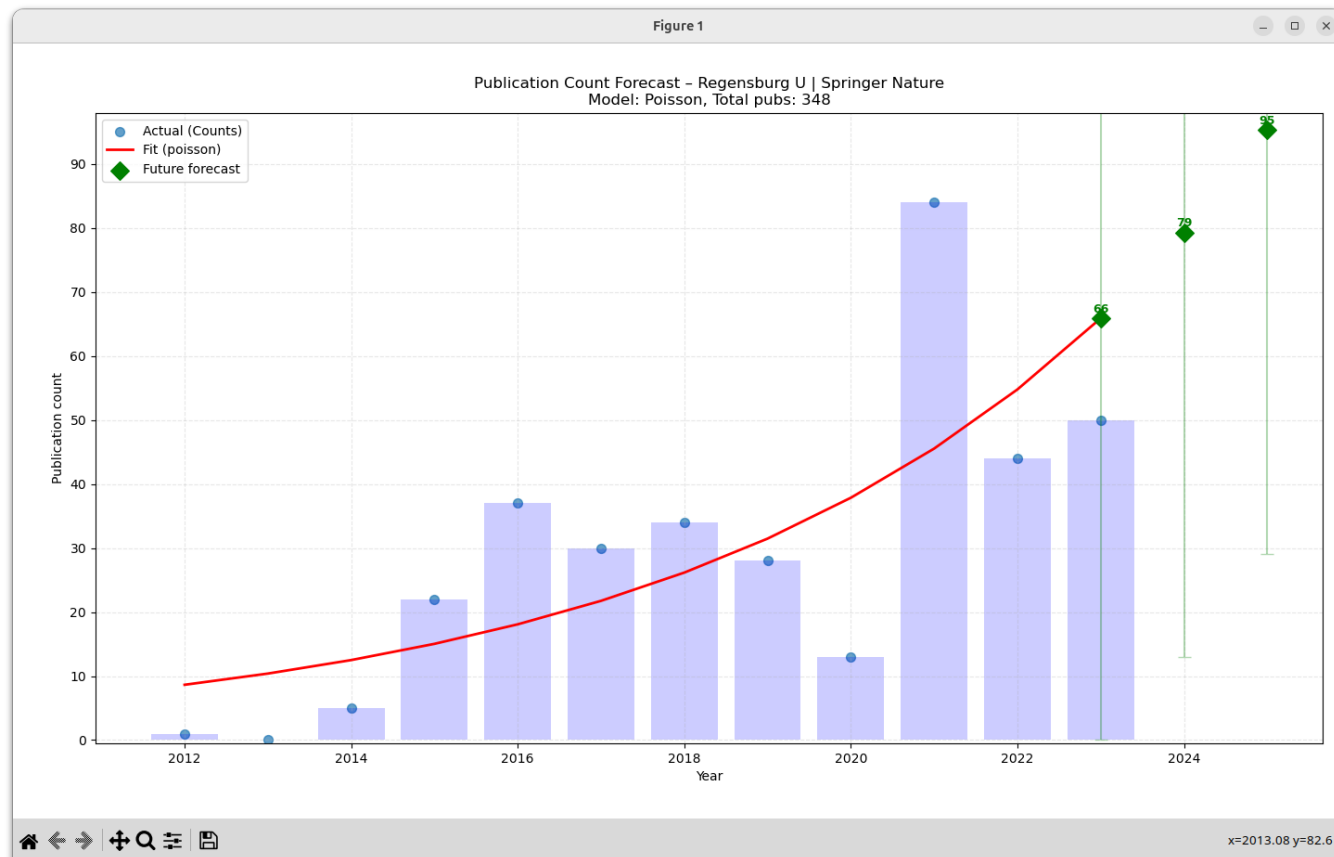


# Übersicht Universität Regensburg





# Vorhersage möglich?





Universität Regensburg

**Dr. Gernot Deinzer**  
Open Access Beauftragter  
Universitätsbibliothek Regensburg

Mehr Daten

**OPENAPC**

# OpenAPC

**250590 Datensätze von**

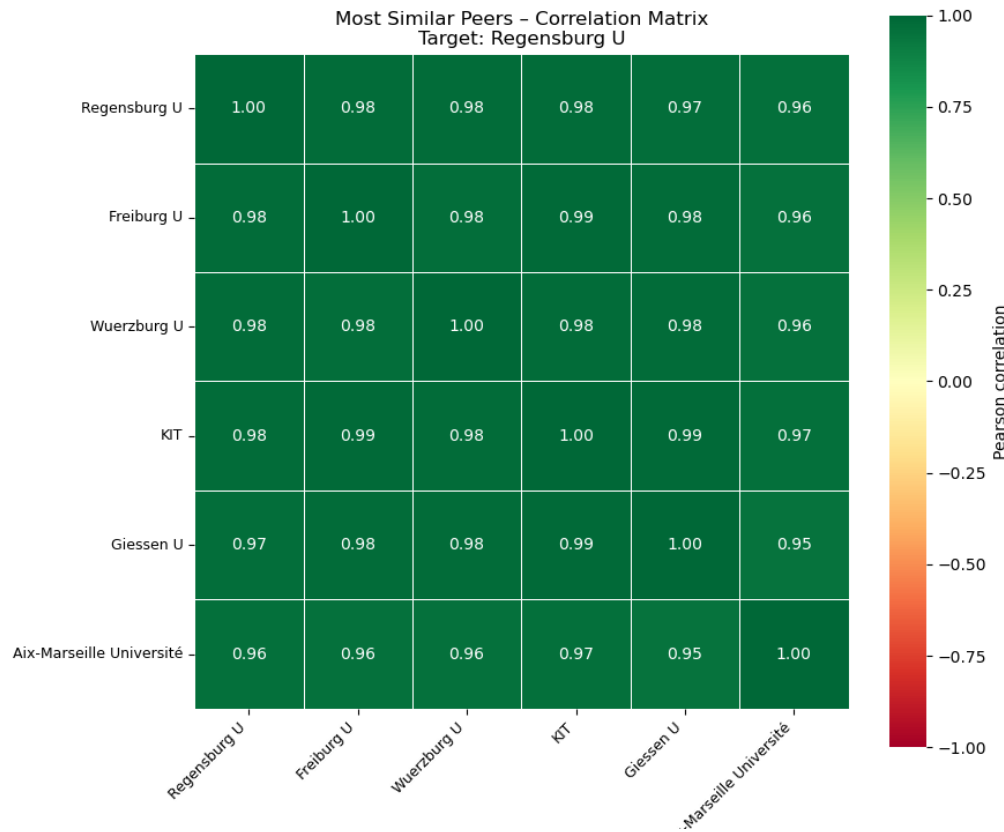
- **469 Institutionen**
- **1079 Verlage**
- **11233 Zeitschriften**

**Maschinelles  
Lernen möglich**

*Wirklich gezahlte Gebühren,  
d.h. inkl. Rabatte etc.*



# Klassische Analyse: Universitäten mit ähnlichen Publikationsverhalten

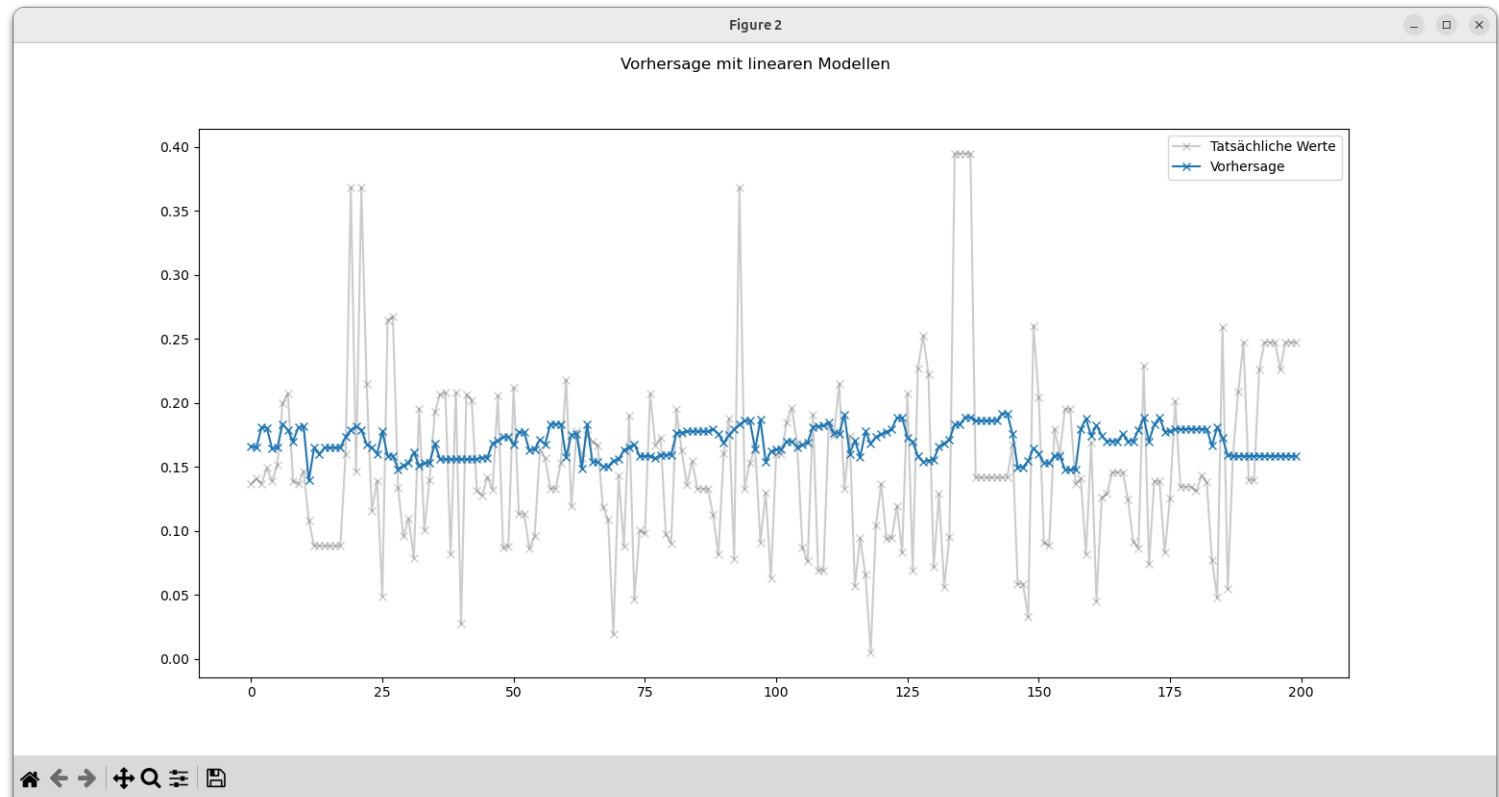


Maschinelles Lernen

# LINEARE REGRESSION

# Lineare Regression

Maximalpreis  
(14.624,20 €) auf 1  
normiert



Kosten

# NICHTLINEARE REGRESSION

# Gradient Boosted Trees

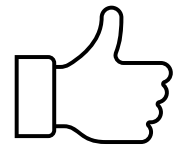
## Kombination aus Regression und Klassifikation

Ausgangslage: schwaches Modell

- > Aufbau auf Entscheidungsbäumen
- > Algorithmus zur Verringerung des Fehlers
- > Optimierung der Verlustfunktion

## Unsere Daten:

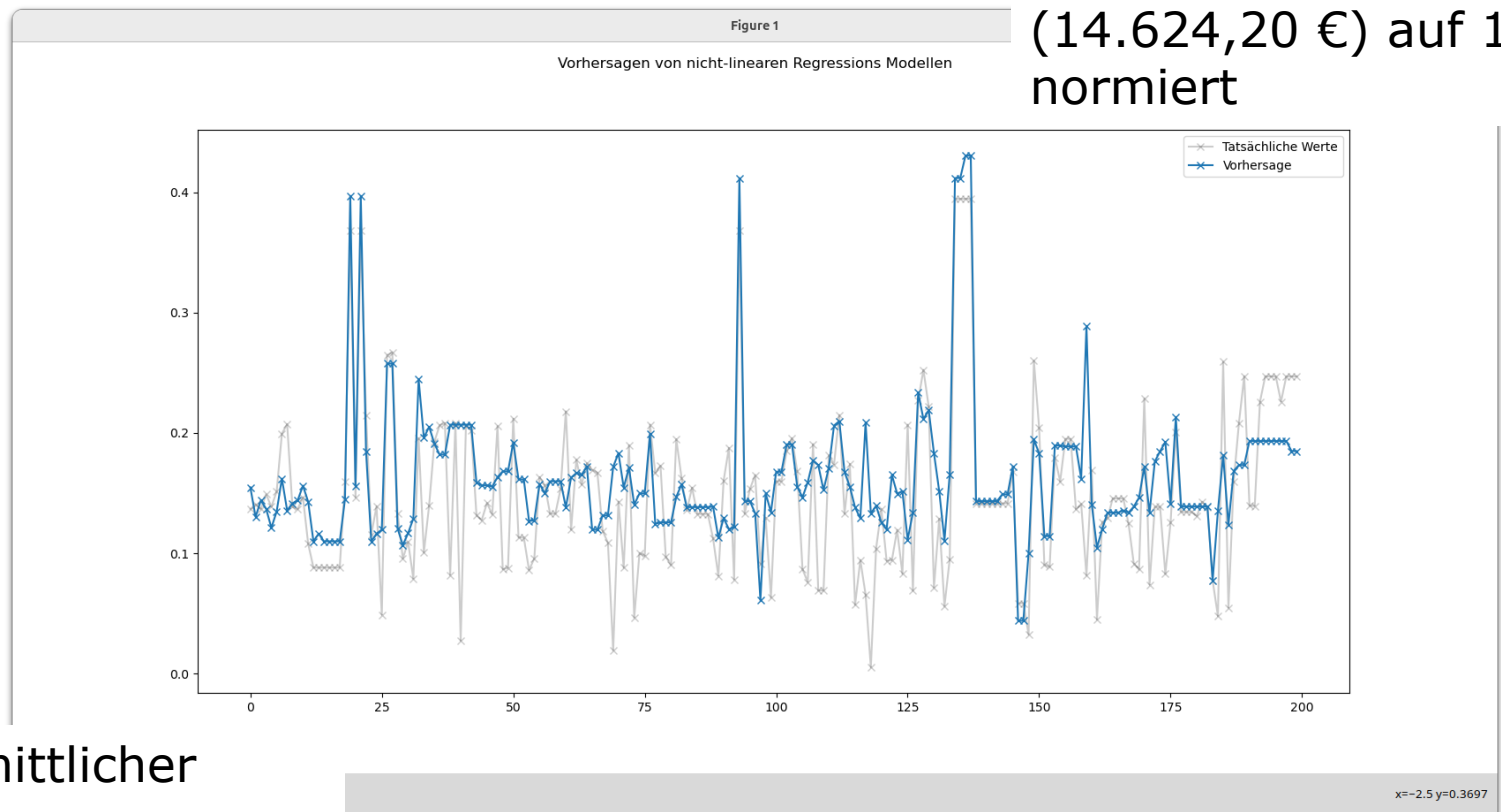
- **Klassifiziert nach Verlag, Zeitschrift und Institution**
- **Ausgewogene Daten**



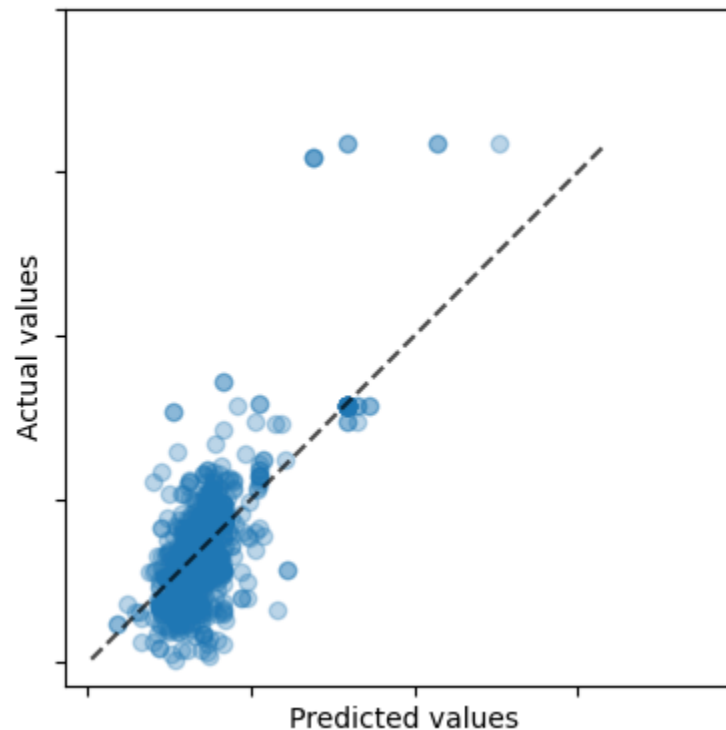


# Vorhersage Publikationsgebühren

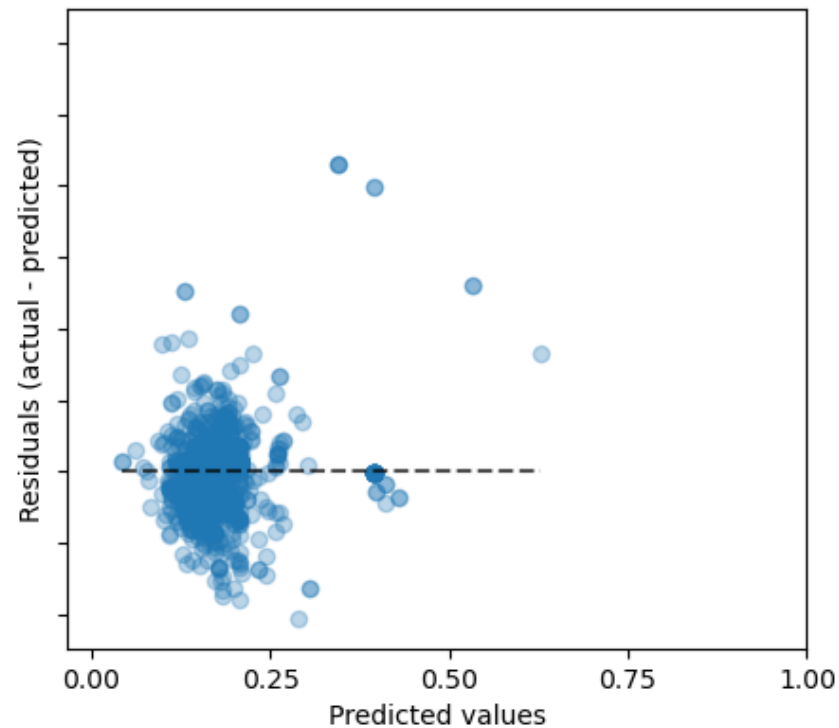
Maximalpreis  
(14.624,20 €) auf 1  
normiert



# Abweichung vs. Vorhersage



# Residuen





Universität Regensburg

**Dr. Gernot Deinzer**  
Open Access Beauftragter  
Universitätsbibliothek Regensburg

# FAZIT

# **Erste Ergebnisse**

- 1. Regensburger Daten nicht ausreichend für maschinelles Lernen**
- 2. Open APC: gute Datengrundlage:**
  1. Verbesserung durch Aktualität
  2. Verbesserung durch größere Teilnahme
- 3. Lineare Regression: nicht geeignet**
- 4. Nichtlineare Regression: gute Ergebnisse**

## **Folgerungen:**

**Maschinelles Lernen und Nichtlineare Regression sind potenzielle Kandidaten für die Vorhersage von APCs**

**OpenAPC in Kombination mit openCost können eine sehr gute und aktuelle Datengrundlage liefern**

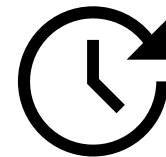
**Neue Herausforderung für Bibliotheken:  
KI-Einsatz bei Planung von Informationsbudget**

**Möglichkeit datenbasierter Information**

# Nächste Schritte

## Vorhersagen für APCs

- Untersuchung: Publikationszahl pro Jahr
- Einschränkung Universität Regensburg
- Abschätzung mittels Berechnung



**Vergleich Vorhersage mit wahren Kosten**



Universität Regensburg

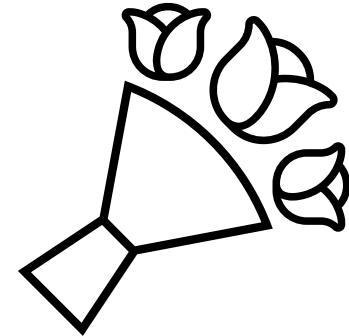
**Dr. Gernot Deinzer**  
Open Access Beauftragter  
Universitätsbibliothek Regensburg

# **DANK UND AUSBLICK**



**Dank**

**openCost**  
**DFG-Projektnummer 457354095**



**Projektteam**

**Open-Access-Publikationskosten**  
**DFG- Projektnummer 491147631**

**Team Publikationsservices – Open Access**

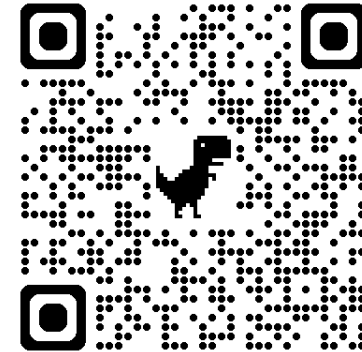
# Navigating Cost Transparency



## openCost Konferenz

**17-19. März 2026**  
**DESY Hamburg**

**Call for Posters**



**Themen (u.a.):**  
**Kosten im Diamond Open Access**  
**Kosten bei Büchern Tagungsbänden etc.**  
**Kosten: internes Format**

