

# PUB – Auswertung von Nutzungsdaten über OA-Statistik (mit einem Ausblick auf OpenAIRE)

Universität Bielefeld
Faculty of Linguistics and Literary Studies




Festwoche: 15. bis 19. Oktober
Portal Studies Staff Research Administration

You are here: Bielefeld University > Faculty of Linguistics & Literary Studies > Staff > Prof. Dr. Schlangen

### David Schlangen's Publications

(This list is automatically generated from the university-wide PUB database; click on [">>>"](#) for details of paper and to download the PDF. If the categorized list doesn't load, click [here](#) for a full list. Some entries may be missing; if you're searching for something you can't find here, again please consult the full list.)

#### Journal Articles

- Schlangen, D., & Skantze, G., 2011. A General, Abstract Model of Incremental Dialogue Processing. *Dialogue and Discourse*, 2(1), 63 - 111.  
**PUB | PDF**
- Rieser, H., & Schlangen, D., 2011. Introduction to the Special Issue of Dialogue and Discourse. *Dialogue and Discourse*, 2(1), 1 - 10.  
**PUB**
- Baumann, T., Buß, O., & Schlangen, D., 2011. Evaluation and Optimization of Incremental Processors. *Dialogue and Discourse*, 2(1), 113 - 141.  
**PUB | PDF**
- Bontas, E.P., & Schlangen, D., 2007. Creating ontologies for content representation - the OntoSeed suite. *Journal on Data Semantics – Special Issue on Extended Papers from 2005 Conferences, IX*.  
**PUB | PDF**
- Schlangen, D., 2005. Modelling dialogue: Challenges and Approaches. *Künstliche Intelligenz*, 3/05, 23 - 28.  
**PUB | PDF**

#### Conference Proceedings

- Schlangen, D., 2012. The Future of Spoken Dialogue Systems is in their Past: Long-Term Adaptive, Conversational Assistants. In *Proceedings of the NAACL-HLT 2012 Workshop on Future Directions and Needs in the Spoken Dialog*

- Prof. Dr. Schlangen
- Professional Activities
- Lehre / Teaching:  
B.A./M.A. Themen;  
Hinweise
- Dialogue Systems Group  
(DSG)
- DSG - People
- Open Positions
- Publications

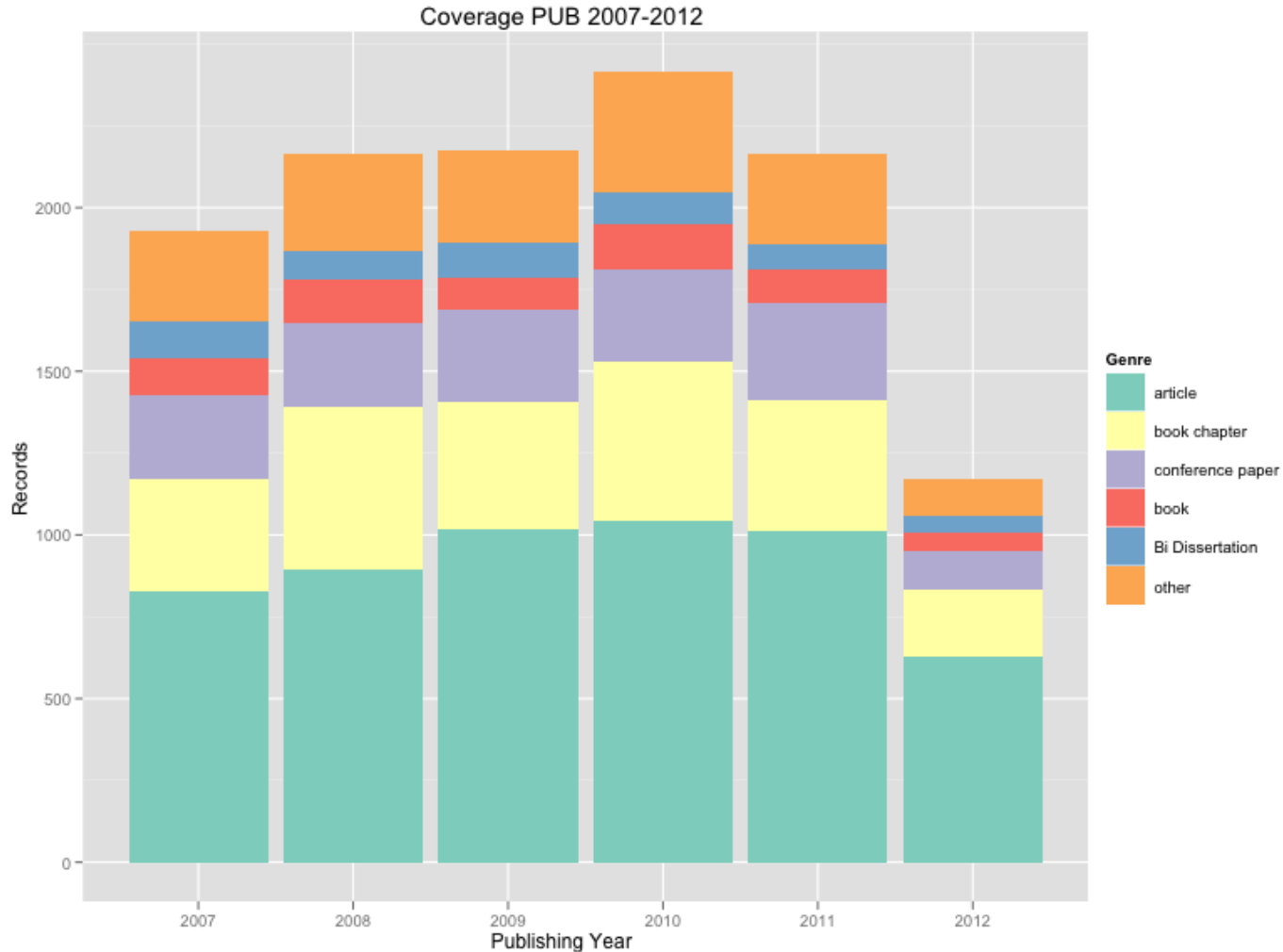
*Jochen Schirrwagen,*  
Universitätsbibliothek  
Bielefeld

OA-Statistik Workshop: Vom Projekt zum Dienst,  
Göttingen, 12.09.2012

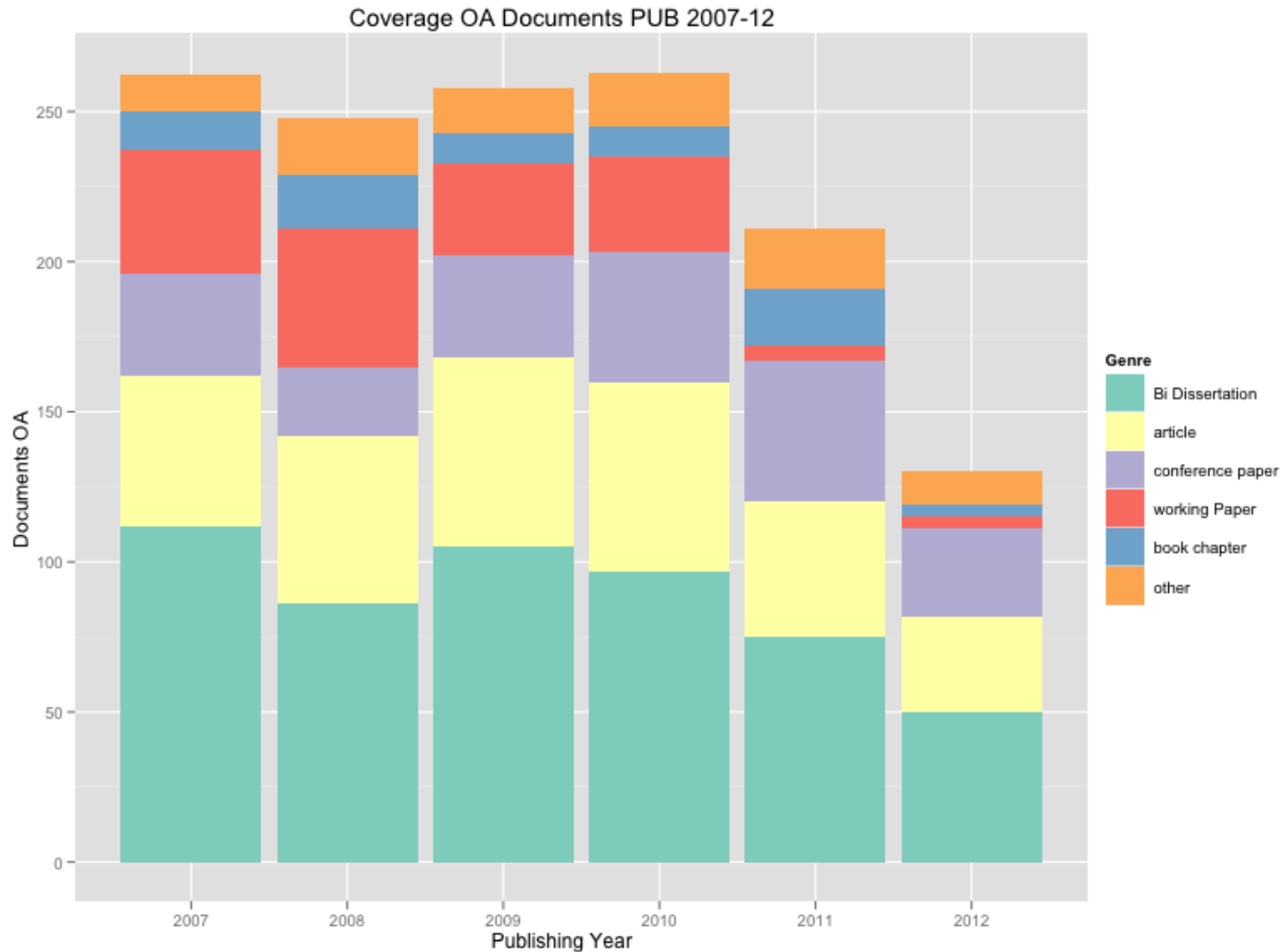
# Publikationen an der Universität Bielefeld

- [Institutionelles Repository](#) für Studierende und Forschende an der Universität Bielefeld
  - Basiert auf [Catmandu/LibreCat](#) Software (Open Source Kollaboration Ghent-Bielefeld-Lund)
  - Nachfolger früherer OPUS-Instanzen (Pre-/Postprints und Dissertationen)
  - OA (5026) und andere (28418 nur Metadaten) Publ.
  - PUB-Anbindung in Einrichtungsportalen
  - Einbettung persönlicher Publikationslisten
- DINI-2010 zertifiziert; registriert in DRIVER / OpenAIRE

# Abdeckung Publikationstypen in PUB



# Abdeckung OA-Publikationstypen in PUB



# Nutzungsdaten – die Schattenseite in der Vergangenheit

- Bisher keine systematische Erhebung von Statistiken
- Zum Teil sind ältere Logdaten verloren gegangen
- Software-Migrationen führten z.T. zu neuen Identifiern für Publ.
- Motivation war gebremst wg. unklarer rechtlicher Situation bzgl.
  - IP-Adressen / Nutzer bezogene Daten
- Wenn statistische Auswertung, dann nur intern
- Fehlende Vergleichbarkeit mit anderen Repositories

# Licht am Horizont

- Nationale Initiativen zu Nutzungsstatistiken + OpenAIRE
  - Entwicklung gemeinsamer [Guidelines \(KE-USG\)](#)
  - Z.B. Format, Transferschnittstelle, Robot-Filterliste
- Motivation von PUB zur Teilnahme an OA-Statistik / OpenAIRE
  - Profitieren von Erfahrungen anderer Teilnehmer
  - Vergleichbarkeit von Statistiken über Repositories hinweg
  - Eine Schnittstelle bedient verschiedene Statistik-Serviceprovider
- ~~(lokale) Alternativen: AWStats / Piwik~~

# Aufsetzen der OA-Statistik Software – Anpassung an lokale Gegebenheiten

- PUB-Konfiguration
  - Landing-Pages, Scripte, CSS unverschlüsselt
  - Volltexte nur über HTTPS
- Jegliche Zugriffe werden über Apache Webserver gespeichert – jedoch
  - getrennte Logfiles für HTTP (combined format) und HTTPS Zugriffe (eigenes Log-Format)

# Anpassungen an die OA-Statistik Software

- „Standard“-Anpassungen in PHP Modulen
  - URL Schema für *abstract, fulltext, any*
  - IP-Klassifizierung für *institutional, administrative*
- Zusätzliche Anpassungen
  - Filterung von Bild-, CSS- und Script-Zugriffen
  - Harvester-Datumsgranularität (PUB liefert nur yyyy-mm-dd)



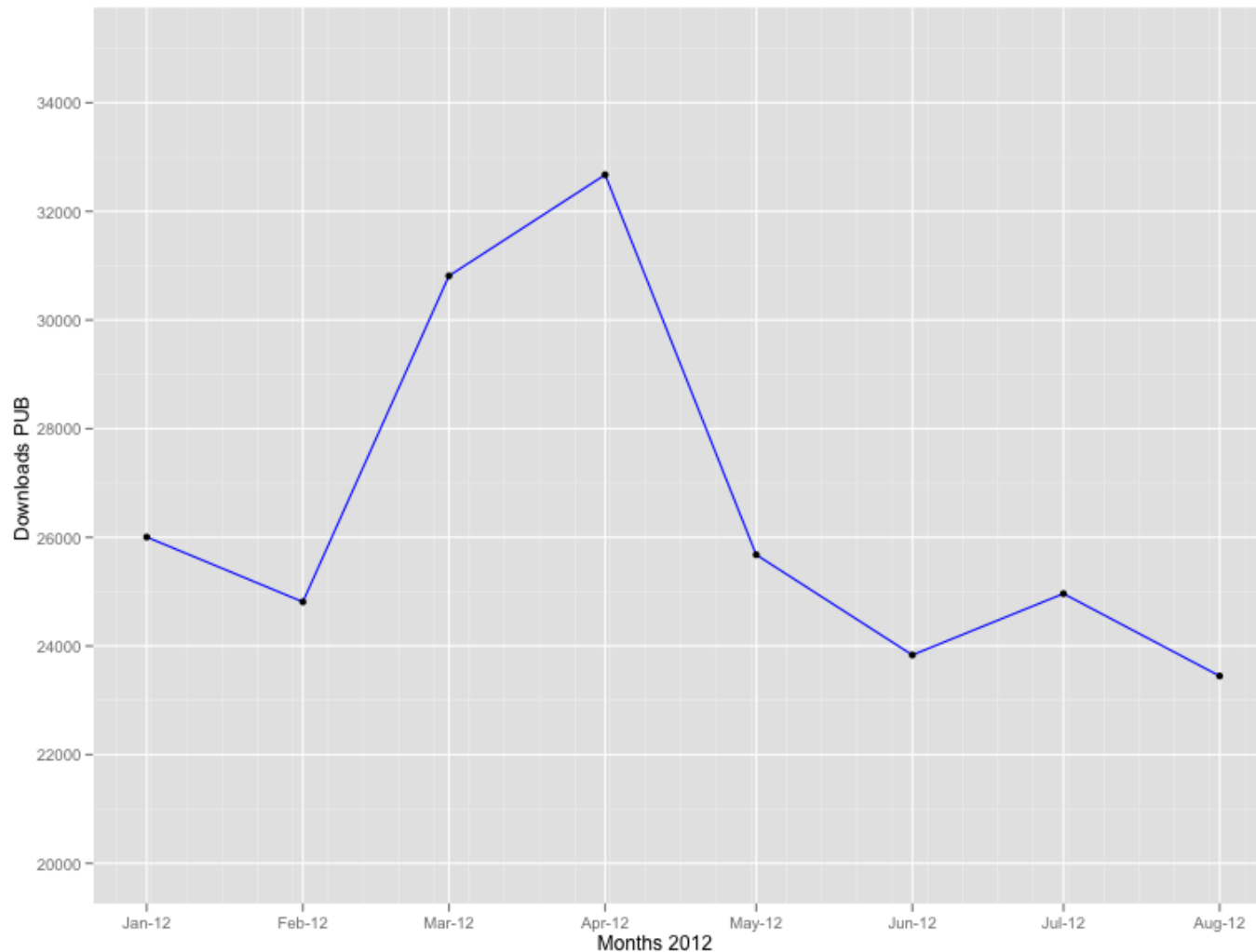
# Optimierungsvorschläge

- Ausführlichere Dokumentation der Software
- Hinweis auf Notwendigkeit zweier Datenbanken
  1. Für OpenURL Context Objects
  2. Für OAI-Identifizier
- Konfiguration ist auf mehrere Dateien verteilt und nicht immer selbsterklärend
- Wozu dient der Request-Typ *any* ?
- REST-API oder FTP für automat. Abruf der Statistiken

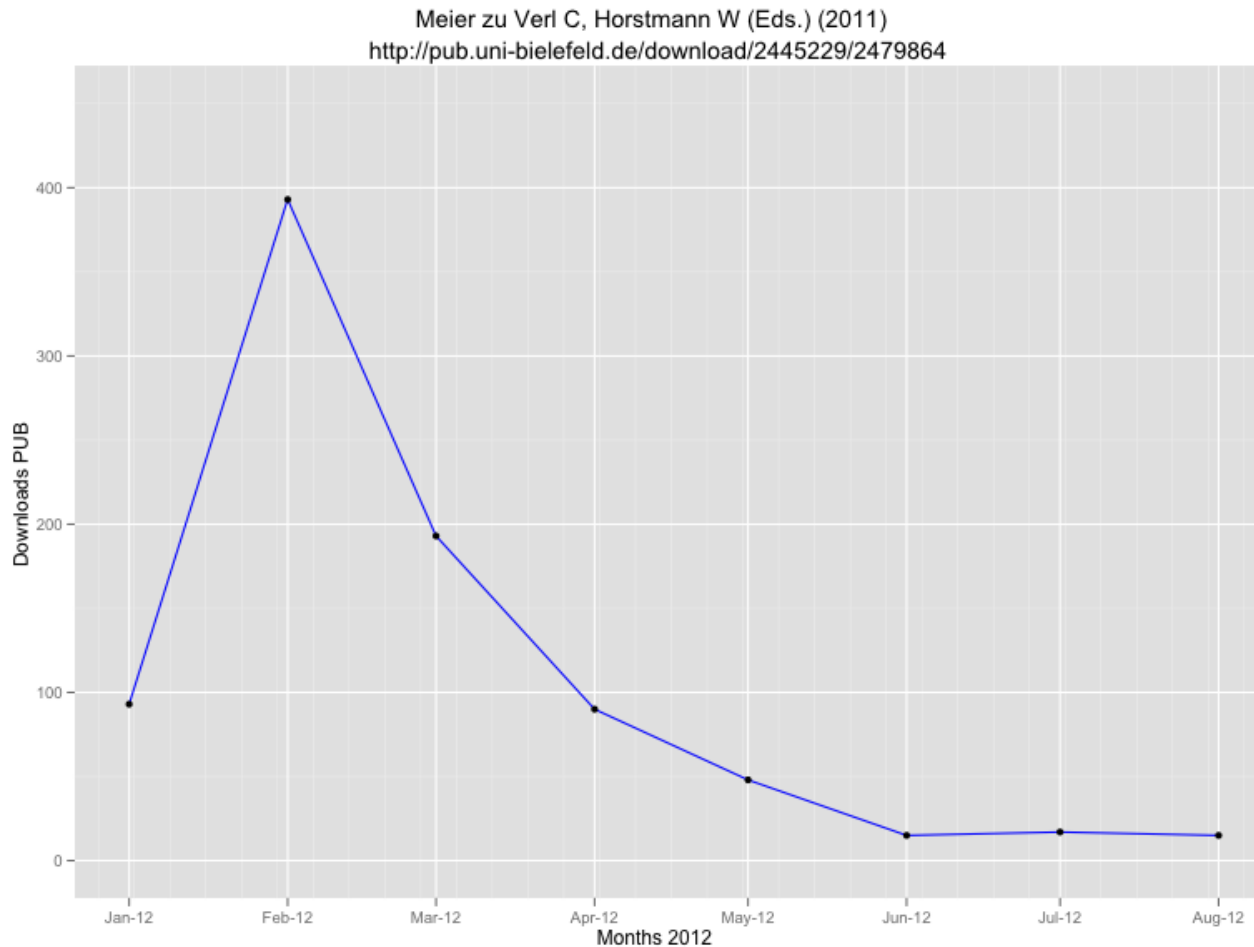
# Erste Ergebnisse aus der Testphase im Sommer 2012

- Statistik wird im JSON Format zurück geliefert
  - Zahl Volltext-Downloads gemäß COUNTER-Standard
  - Zahl Metadata-Views
- Manuelles Prüfen auf Plausibilität und Rückmelden evtl. Probleme / Auffälligkeiten
- Nutzungszahlen Primärpublikationen > Nutzungszahlen Metadatennachweise
  - U.a. abh. vom Publikationstyp (z.B. Dominanz von Dissertationen)

# PUB Volltext-Downloads Jan-Aug 2012 gesamt



# Volltext-Download Charakteristiken einer ausgewählten Publikation



# Einbettung der Statistiken in PUB - Pläne

- Einbettung zunächst nur für OA-Publikationen angedacht
- Grafische Darstellung möglichst unter Berücksichtigung des dynamischen Verlaufs der Nutzungszahlen
- Berücksichtigung von Statistiken anderer Serviceprovider soweit möglich, z.B. PMC-Zahlen, PLOS-ALM Zahlen
- Übergreifende Vergleichsmöglichkeit nach Publikationstyp, -jahr

# Ausblick auf Nutzungsstatistiken in OpenAIRE - Ansatz

- Ausgangspunkt: Publikationen in FP7 SC39 Projekten  
nachgewiesene FP7 Publ. gesamt: ~27.000
- Abschätzung des Publikationsoutputs von FP7  
Projekten
- Aufbau eines [internationalen Netzwerks von  
Nutzungsdatenanbietern](#)
- Grafische (dynamische) Visualisierung und Reporting
- Komplementärer Ansatz:  
Bibliometrie, Webometrie, Nutzungsstatistiken  
(auch bedingt durch Randbedingungen)

# Ausblick auf Nutzungsstatistiken in OpenAIRE - Fragestellungen

- Verknüpfung Nutzungsdaten mit alternativen Metriken  
z.B. PLOS-ALM

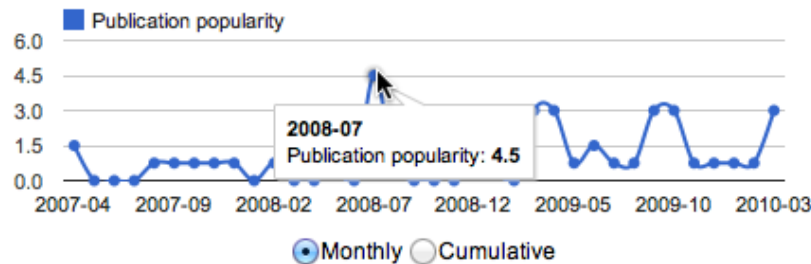
## Fragestellungen

- Was sind geeignete Indikatoren zur Messung der Nutzung von Projektergebnissen?
- Wie lassen sich unterschiedliche Metriken kombinieren?
- Wie sind die Statistiken zu interpretieren?

# Nutzungsstatistiken in OpenAIRE – Beispiel Visualisierung Popularity-Indikator

## Article- Level

Key management and link-layer security of wireless sensor networks : Energy-efficient attack and defense  
Author(s): Law, Y.W.



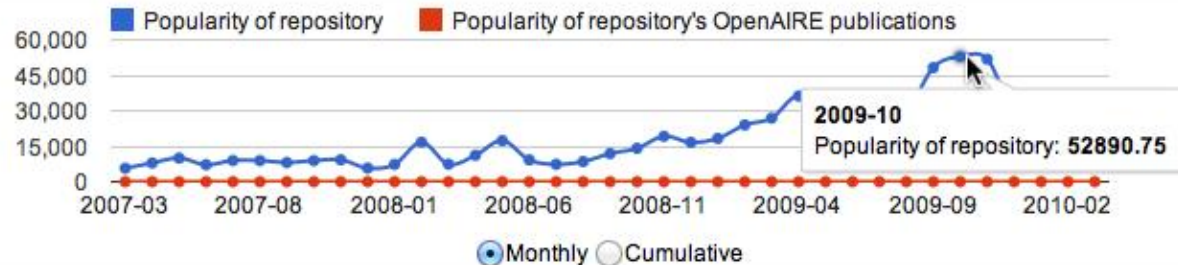
Details ↓ BibTeX ↓ Statistics ↑



## Repository- Level

Universiteit Twente Repository

Show details ↓ Hide statistics ↑





# Danke für Ihre Aufmerksamkeit!

## Kontakt:

- Jochen Schirrwagen  
<jochen.schirrwagen AT uni-bielefeld.de>
- Mathias Lösch  
<mathias.loesch AT uni-bielefeld.de>
- Najko Jahn  
<najko.jahn AT uni-bielefeld.de>