Scholarly Research and Publications Over Time: Identifying Trends for an Open Access Journal by Applying Data-Mining Methods

Judith Schoßböck, Danube University Krems
Noella Edelmann, Danube University Krems
Nina Rizun, Gdansk University of Technology
Anneke Zuiderwijk, Delft University of Technology

DOI: https://doi.org/10.7557/5.5606
Munin Conference 2020, 18.11.2020

Copyright (c) 2020 Judith Schoßböck, Noella Edelmann, Nina Rizun, Anneke Zuiderwijk
This work is licensed under a Creative Commons Attribution 4.0 International License.
JeDEM.org: Journal Management & Editorship

Judith Schoßböck & Noella Edelmann
Journal of eDemocracy and Open Government
Scholarly-led & interdisciplinary community, platinum OA
Danube University Krems
Data Perspective

Gdansk University of Technology
Data mining & semantic analysis

Delft University of Technology
Open data, open science, open government
Editor in Chief JeDEM
"Vital Statistics" of JeDEM (Edgar & Willinsky, 2010)

<table>
<thead>
<tr>
<th>Journal basics &amp; editorial practices</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Journal sponsor</td>
<td>Academic department</td>
</tr>
<tr>
<td>Acceptance rate</td>
<td>50%</td>
</tr>
<tr>
<td>Research areas/diversity</td>
<td>E-democracy, e-society, e-participation, e-government, open data, data sharing &amp; use, open science, open access, open source</td>
</tr>
<tr>
<td>Publication</td>
<td>Since 2009 (2x/year)</td>
</tr>
<tr>
<td>Indexing</td>
<td>Scopus, EBSCO, DOAJ, Google scholar, Public knowledge project (PKP) metadata harvester; DOIs</td>
</tr>
</tbody>
</table>
User Survey 2019, N=57 (Reasons for publishing with JeDEM)

2.5 JeDEM publishes high quality articles

2.6 JeDEM has a good reputation in the field

2.7 JeDEM is an e-journal

2.8 JeDEM is open access

2.11 JeDEM is indexed (Scopus, EBSCO, DOAJ, Google scholar and the Public Knowledge Project metadata harvester)
Interdisciplinary Community (reasons for publishing with JeDEM)

affinity to research area

value-driven: OA

wide distribution (numbers)
good experience :)
support of a younger journal

turnaround time

stable and recognized publisher vs commercial journals

high academic standard/quality

simple submission process

indexing & credits/including criticism
Problem Statement

**Strategic choices** of journal editors & managers
- Enhance visibility, accessibility, impact
- Support of authors with publication choice

Useful information for **authors**
- Metrics and quantitative journal measures
- Research areas, themes and trends

[Logos for EBSCO, Scopus, Google Scholar, and DOAJ]
Objective

Approach for journal managers to identify research trends, readers’ preferences & develop a digital mindset

Derive factors that impact trends, cooperations & journal development over time

WHERE?
- Countries & research cooperations

WHAT?
- Topics of papers over time
Methods

Data mining of journal content and metadata of all articles published since 2009

- text mining, topic modelling, k-means clustering, social network analysis, community detection

Qualitative Interpretation

- internal journal management perspective
- relate trends to strategic decisions or publication details
Results

1. Publication quantity and types
2. Leading countries
3. Visibility of research cooperations (through publications)
4. Research topics and evolution over time
5. Research topics per countries
6. Sub-topics of papers over time
7. Co-occurrence of words over time
1. Publication Quantity and Types

- **RESEARCH PAPERS**
- **EDITORIAL**
- **INVITED PAPERS**
- **CASE STUDIES**
- **SPECIAL ISSUE: REFLECTIONS**
- **SPECIAL ISSUE: PROJECT DESCRIPTIONS**

The chart shows the quantity and types of publications from 2009 to 2020, with different colors representing different types of publications.
2.1. Countries
2.1. Countries (Europe)
3. Visibility of Research Cooperations (Publications)
4. Research Topics (Methodological Steps)

• Preprocessing of article title & abstract
• Separation into epochs ordered over time
• Extracting the 10 most frequently used words and bigrams
• Relating the keywords (and bigrams) characterizing the main thematic context of each epoch with the parameters of adjacent eras in time
• Frequency matrix
• Cosine similarity
• K-means clustering
4.1. Research Topics (Clusters)

3 clusters of thematic streams (bigrams)

(1) ELECTRONIC PETITIONS, E-PARTICIPATION INITIATIVE & CITIZEN PARTICIPATION

(2) OPEN DATA, SOCIAL MEDIA & SOCIAL NETWORKING, DISRUPTIVE TECHNOLOGIES & DIGITAL TRANSFORMATION

(3) OPEN DATA-DRIVEN GOVERNMENT, SMART CITY & CO-CREATION
4.2. Research Topics (Topical Trend)
5.1. Research Topics per Continents

- **Europe**: Topic 1. Electronic petitions, e-participation initiative & citizen participation
- **North America**: Topic 2. Open data, social media & social networking, digital transformation & disruptive technologies
- **Asia**: Topic 3. Open data-driven government, smart city & co-creation
- **South America**: Topic 2. Open data, social media & social networking, digital transformation & disruptive technologies
- **Australia/Oceania**: Topic 2. Open data, social media & social networking, digital transformation & disruptive technologies
- **Africa**: Topic 3. Open data-driven government, smart city & co-creation
### 6. Sub-topics of Papers (Topic Modelling)

<table>
<thead>
<tr>
<th>Topic 1. Electronic petitions, e-participation initiative &amp; citizen participation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1. Online Deliberation and Electronic Support</td>
</tr>
<tr>
<td>1.2. Technology and E-Petition Systems</td>
</tr>
<tr>
<td>1.3. Citizens e-Participation Initiatives</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Topic 2. Open data, social media &amp; social networking, digital transformation &amp; disruptive technologies</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1. Disruptive Technologies and Political Participation</td>
</tr>
<tr>
<td>2.2. Open Data and Open Government</td>
</tr>
<tr>
<td>2.3. Social Networking and Social Media Planforms</td>
</tr>
<tr>
<td>2.4. Decision-making and Digital Transformation</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Topic 3. Open data-driven government, smart city &amp; co-creation</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1. Open Linked Data and Smart City</td>
</tr>
<tr>
<td>3.2. Open Data-Driven Government and Civic Technologies in Public Sector</td>
</tr>
<tr>
<td>3.3. Digital Political Participation and Citizen Engagement</td>
</tr>
<tr>
<td>3.4. Electronic Voting</td>
</tr>
</tbody>
</table>
7. Word Co-occurrence Over Time
Work in Progress

- Research methodologies over time
- Primary data collection methods
Influential Factors of Journal Development

- **Guest editors** → topical and regional influence
- **Indexing systems** → quantity & quality of submissions
- **Research cooperations** → topical, regional, quantity influence
- **Naming/strategic name change (“Open Government”)** → reflecting research trends → topical influence
- **Predominance/missing of countries** → Hegemonial bias? (f.i. Africa & e-participation)
- **Other drivers** of topical & journal development?
Discussion

1. Tracing the evolutionary stages of the study of scientific problems & critical assessment of scholarly developments
   ✓ Are indexing systems trumping topical development?
   ✓ Who takes part in shaping research trends?
   ✓ Which countries are missing?

2. Insights regarding journal development & strategy
   ✓ Do our calls match what we get submitted? What difficulties are related to being interdisciplinary?
   ✓ Usefulness for other journal editors and audiences?
THANK YOU!

Case Comparison?
Similar analysis?

Contact:
www.jedem.org
judith.schossboeck@donau-uni.ac.at