B!SON
– algorithmenbasiertes Empfehlungssystem für Open-Access-Zeitschriften

Anita Eppelin, https://orcid.org/0000-0003-3184-5930

Elias Entrup, https://orcid.org/0000-0002-7380-1189

Kolloquium Wissensinfrastrukturen, Universitätsbibliothek Bielefeld

16. Dezember 2022

B!SON
Bibliometric and Semantic Open Access
Recommender Network

Agenda

1. Fragestellung und Idee
2. Vorgehen
3. Betaversion
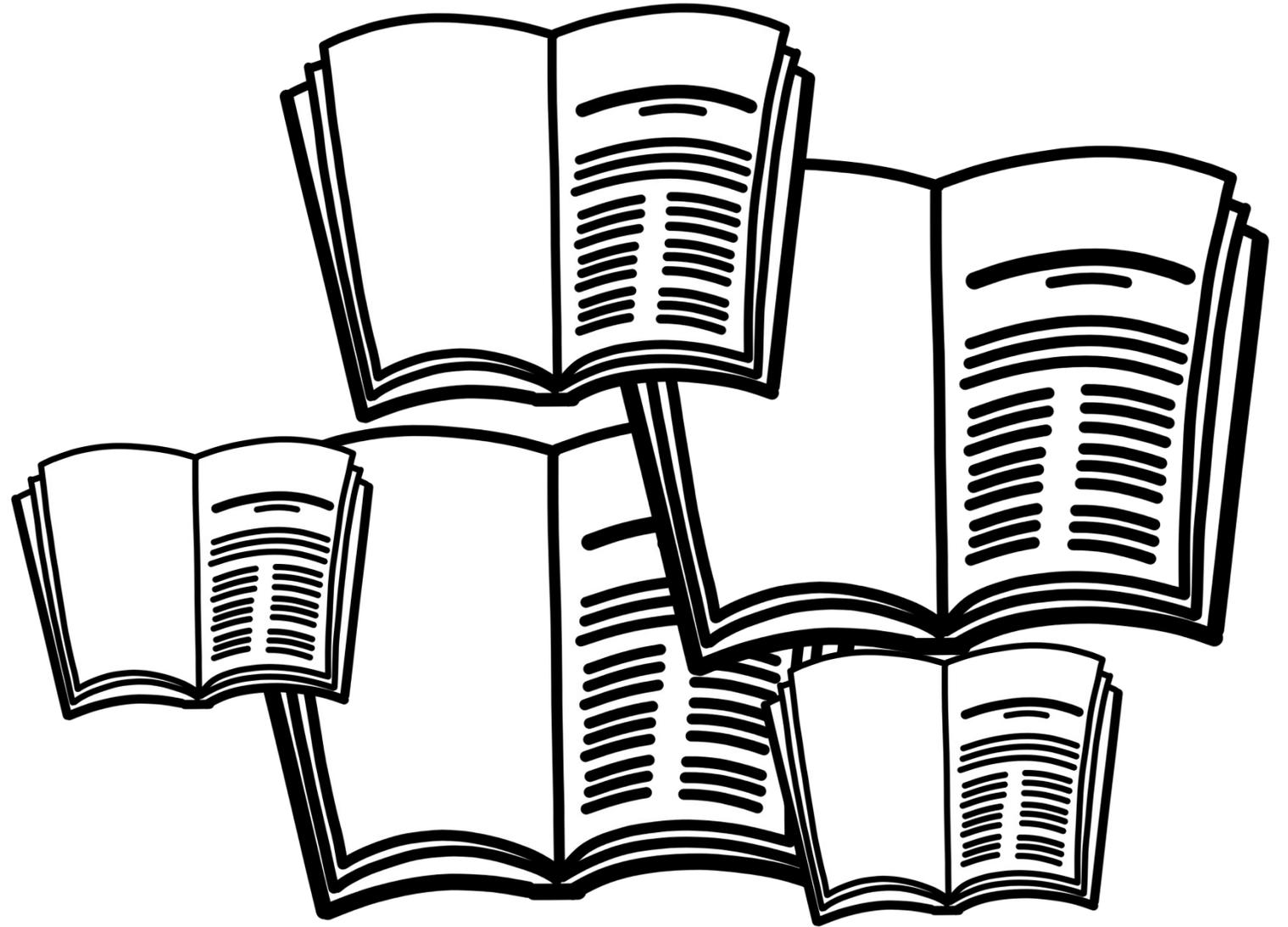4. Technische Verfahren hinter den Empfehlungen
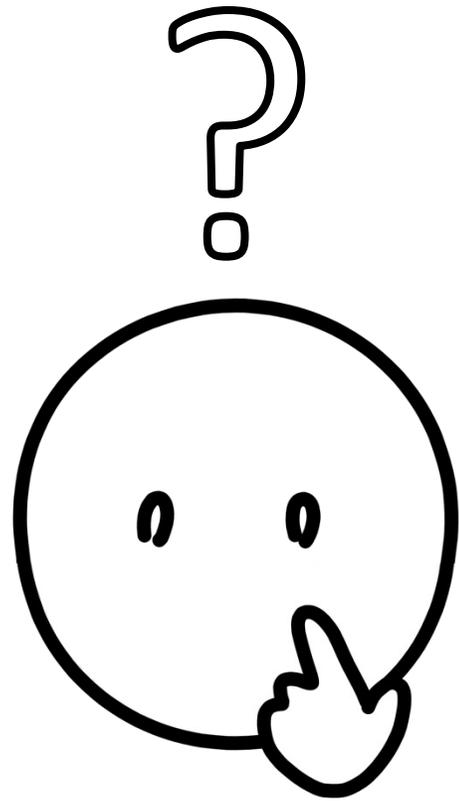5. Lokale Einbindung
6. Ergebnisse und Ausblick
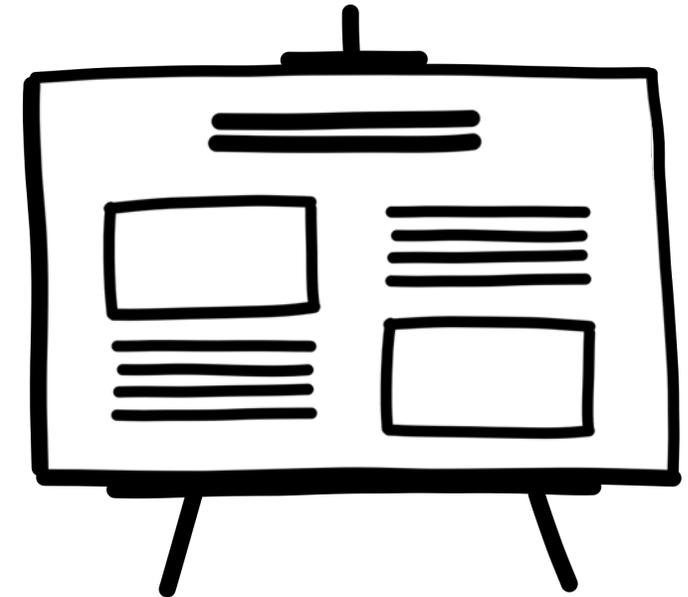
B!SON

1. Fragestellung und Idee

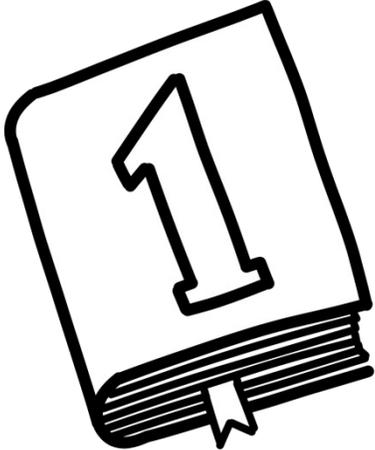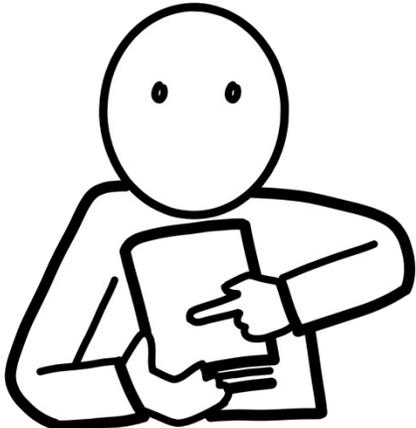Happy researcher has a nice article for the science world …

… and lots of journals to publish in

A lot to consider
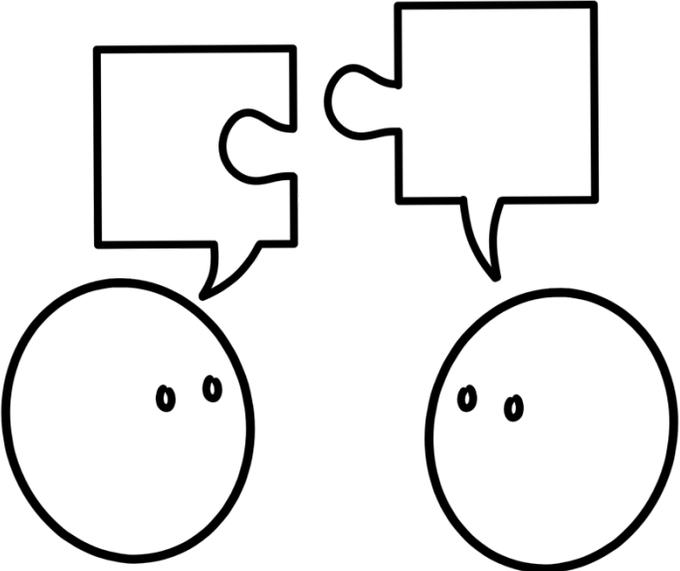
A lot to consider

B!SON to the rescue!

1. journal X (medium-size)

2. journal Y (big)

3. journal Z (small)
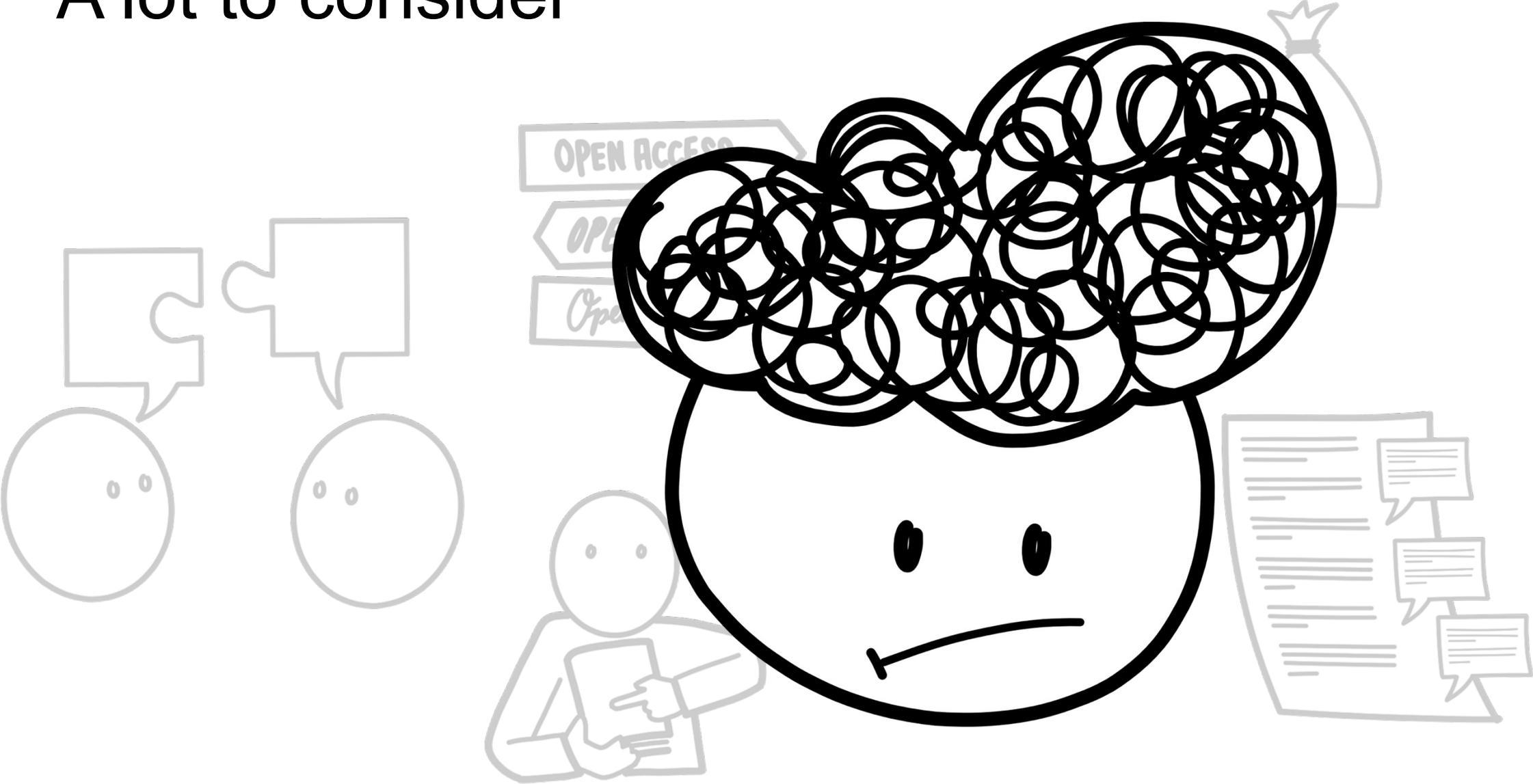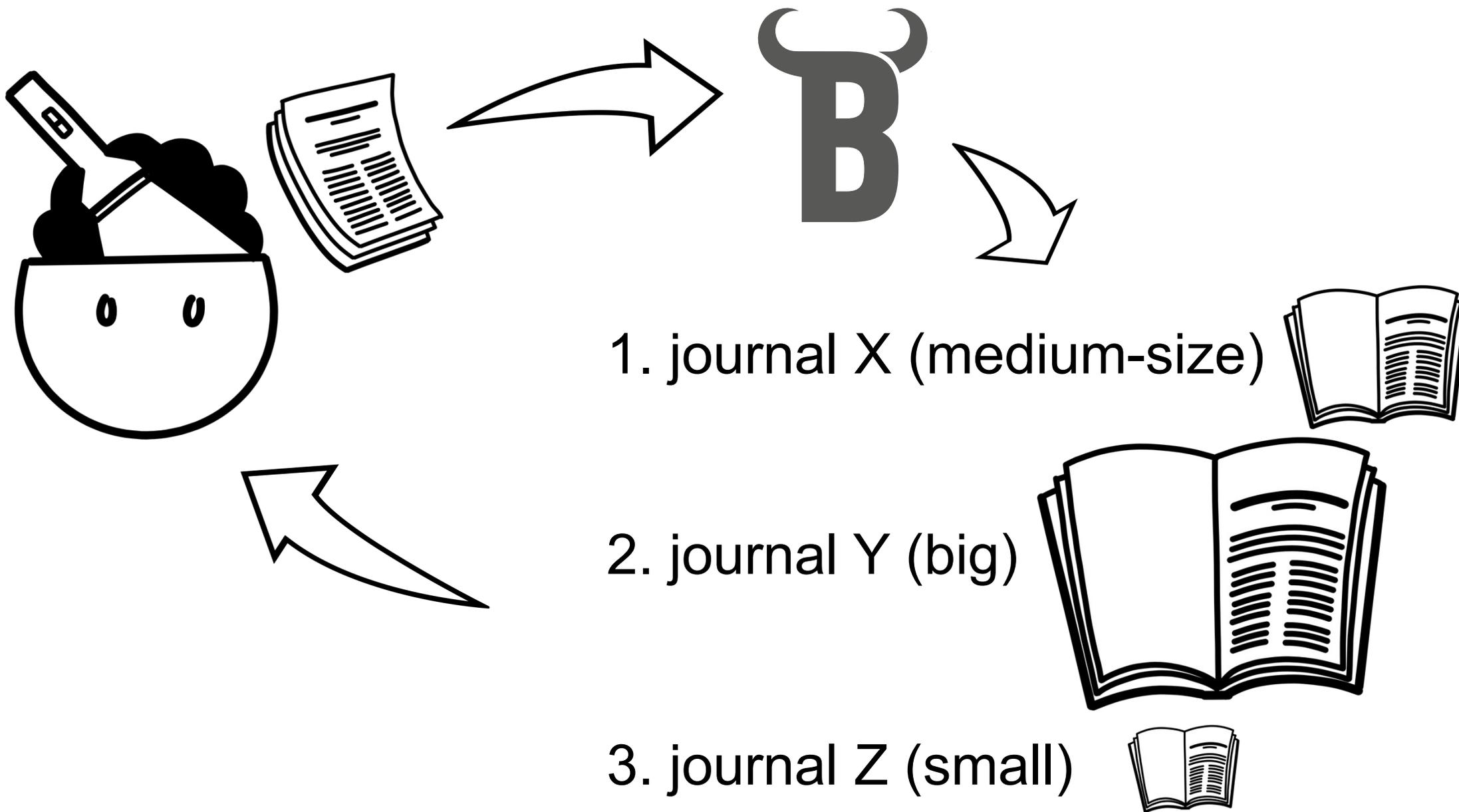
B!SON empfiehlt fachlich relevante, qualitätsgesicherte Open-Access-Zeitschriften

– open source, verlagsagnostisch, datenschutzkonform, kostenfrei

B!SON

2. Vorgehen

Zeitstrahl

Q2 2021
- Anforderungsanalyse
- Community-Aufbau

Q3 2021
- Analyse der Datenbasis (DOAJ, OC)

Q4 2021
- Entwicklung Empfehlungsalgorithmus:
  semantische Verfahren
- Bibliometrische Verfahren:
  Tests und erste Implementierung

Q1 2022
- B!SON Web Interface und API

Photo: James Brooks, CC BY 2.0, via Wikimedia Commons

Zeitstrahl

Q2 2022
- Launch Betaversion
- Community-basierte Qualitätstests

Q3 2022
- Erweiterung bibliometrischer Verfahren

Q4 2022
- **Entwicklung lokale Anpassung/Einbindung**
- **Nachhaltigkeits-/Verstetigungskonzept**

Q1 2023
- **Launch Produktivversion**

Photo: James Brooks, CC BY 2.0, via Wikimedia Commons
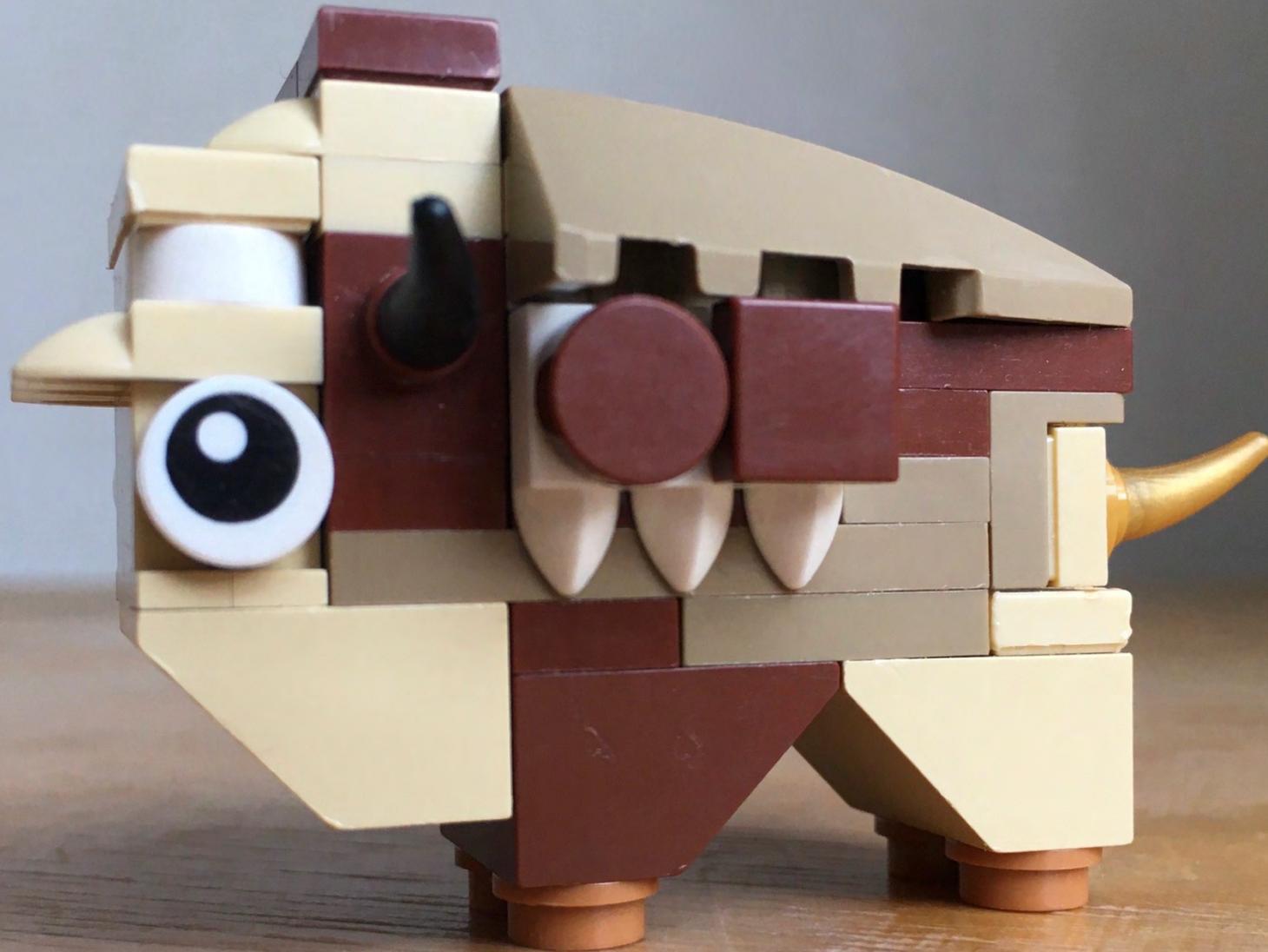
3. B!SON-Betaversion

B!SON helps you to find a suitable Open Access journal for your publication by leveraging semantic and bibliometric methods. Simply enter the details of your manuscript below or fetch the details of a paper via its DOI or arXiv ID. It is not necessary to provide all fields. The journal information is provided by the DOAJ. We cannot guarantee the correctness and details (like costs) should be checked on the journal homepage before submitting.

Title

    Title of your paper...

Abstract

    Abstract of your paper...

References ?

    References of your paper...

Search   Clear

Give us feedback

TIB   SLUB Wir führen Wissen.   Bundesministerium für Bildung und Forschung   DOAJ   OC   @oa_bison

B!SON helps you to find a suitable Open Access journal for your publication by leveraging semantic and bibliometric methods. Simply enter the details of your manuscript below or fetch the details of a paper via its DOI or arXiv ID. It is not necessary to provide all fields. The journal information is provided by the DOAJ. We cannot guarantee the correctness and details (like costs) should be checked on the journal homepage before submitting.

Title

ORCID-linked labeled data for evaluating author name disambiguation at scale

Abstract

How can we evaluate the performance of a disambiguation method implemented on big bibliographic data? This study suggests that the open researcher profile system, ORCID, can be used as an authority source to label name instances at scale. This study demonstrates the potential by evaluating the disambiguation performances of Author-ity2009 (which algorithmically disambiguates author names in MEDLINE) using 3 million name instances that are automatically labeled through linkage to 5 million ORCID researcher profiles. Results show that although ORCID-linked labeled data do not effectively represent the population of name

References ❓

10.1371/journal.pone.0158731
10.1002/asi.22621
10.1002/asi.23063
10.1080/10572317.2016.1243962
10.1016/j.joi.2013.06.006
10.1371/journal.pone.0195773

[Search] [Clear]

100 Results

[Filter] [Export] [List | Table] [Score ⌄]

According to your input, we automatically set the language filter to "English". You can change it along with the other filters. Setting the subject filter to "Bibliography. Library science. Information resources" will likely result in more specific results. ✕

100 Results

Filter | Export | List | Table | Score ⌄

Score: 100% ▼

Quantitative Science Studies

Publisher: The MIT Press

Keywords: scholarly communication, science indicators, science policy, scientific workforce, system of science

Score: 88% ▼

IEEE Access

Publisher: IEEE

Keywords: communications, engineered materials, engineering, manufacturing

Score: 74% ▼

JLIS.it

Publisher: Firenze University Press

Keywords: archival science, librarianship, library science

Score: 25% ▼

Humanities & Social Sciences Communications

Publisher: Springer Nature

Keywords: behavioural sciences, humanities, social sciences

Score: 25% ▼

PLoS ONE

Publisher: Public Library of Science (PLoS)

100 Results

According to your input, we automatically set the language filter to "English". You can change it along with the other filters. Setting the subject filter to "Bibliography. Library science. Information resources" will likely result in more specific results.   ✕

Score: 100%

Quantitative Science Studies

Publisher: The MIT Press

Keywords: scholarly communication, science indicators, science policy, scientific workforce, system of science

Score: 88% ▼

IEEE Access

Publisher: IEEE

Keywords: communications, engineered materials, engineering, manufacturing

Score: 74% ▼

JLIS.it

Publisher: Firenze University Press

Keywords: archival science, librarianship, library science

Score: 25% ▼

Humanities & Social Sciences Communications

Publisher: Springer Nature

Keywords: behavioural sciences, humanities, social sciences

Score: 25% ▼

PLoS ONE

Publisher: Public Library of Science (PLoS)

Filter   Export   List | Table   Score ▼

According to your input, we automatically set the language filter to "English". You can change it along with the other filters. Settin...   ...pecific results. ✕

Recommendation based on these similar articles

Similarities in title:
- Author name disambiguation of bibliometric data: A comparison of several unsupervised approaches
- Analyzing academic mobility of U.S. professors based on ORCID data and the Carnegie Classification

Similarities in abstract:
- Author name disambiguation of bibliometric data: A comparison of several unsupervised approaches
- Analyzing academic mobility of U.S. professors based on ORCID data and the Carnegie Classification

Articles citing the references:
- 10.1162/qss_a_00081

The neural network component suggested this journal with a confidence compared to other journals of: 69%

Score: 100%

Score: 88% ▼

...n of science

Keywords: communications, engineered materials, engineering, manufacturing

Score: 74% ▼

JLIS.it

Publisher: Firenze University Press

Keywords: archival science, librarianship, library science

Score: 25% ▼

Humanities & Social Sciences Communications

Publisher: Springer Nature

Keywords: behavioural sciences, humanities, social sciences

Score: 25% ▼

PLoS ONE

Publisher: Public Library of Science (PLoS)

100 Results

Filter | Export | List Table | Score ⌄

Score: 100%

Quantitative Science Studies

Publisher: The MIT Press

Keywords: scholarly communication, science indicators, science policy, scientific workforce, system of science

Score: 88% ▼

IEEE Access

Publisher: IEEE

Keywords: communications, engineered materials, engineering, manufacturing

Score: 74% ▼

JLIS.it

Publisher: Firenze University Press

Keywords: archival science, librarianship, library science

Score: 25% ▼

Humanities & Social Sciences Communications

Publisher: Springer Nature

Keywords: behavioural sciences, humanities, social sciences

Score: 25% ▼

PLoS ONE

Publisher: Public Library of Science (PLoS)

100 Results

Journal	Score	Alternative title	Publisher	Maximum publication fees (APCs)	Language(s)	Author retains unrestricted rights	Article receives DOI	P n
Quantitative Science Studies	**100%**	-	The MIT Press	800 USD (approx. 761.9€)	EN	✓	✓	B
IEEE Access	**88%**	-	IEEE	1850 USD (approx. 1761.9€)	EN	✓	✓	B
JLIS.it	**74%**	Italian Journal of Library and Information Science	Firenze University Press	-	EN, IT	✓	✓	C r
Humanities & Social Sciences Communications	**25%**	-	Springer Nature	990 GBP (approx. 1152.5€)	EN	✓	✓	C r
PLoS ONE	**25%**	PLOS ONE	Public Library of Science (PLoS)	1805 USD (approx. 1719.05€)	EN	✓	✓	B
Problemos	**23%**	-	Vilnius University Press	-	EN, LT	✓	✓	C r
EPJ Data Science	**23%**	-	SpringerOpen	1190 GBP (approx. 1385.33€)	EN	✓	✓	B
BMC Bioinformatics	**21%**	-	BMC	1690 GBP (approx. 1967.4€)	EN	✓	✓	B

☐ Author retains unrestricted copyrights and publishing rights

Select keywords ❓

[Search keywords_____] [✕]

Average publication time: 52 weeks

●───────────────────────────────────────●

☐ academic libraries
☐ advertising discourse
☐ advocacy
☐ allergology
☐ amputation
☐ archaeology

Maximum publication fees (APCs): 5100 €

●───────────────────────────────────────●

☐ archival science
☐ archival studies
☐ archive studies
☐ arrhythmia

Select subject ❓

[All   ⌄]

☐ art
☐ artificial nutrition

Select language

[English   ⌄]

[✖ Reset filters]

According to your input, we automatically set the language filter to "English". You can change it along with the other filters. Setting the subject filter to "Bibliography. Library science. Information resources" will likely result in more specific results.   ✕

Score: 100% ▼

Quantitative Science Studies

Publisher: The MIT Press

Keywords: scholarly communication, science indicators, science policy, scientific workforce, system of science

100 Results

Filter Export | List | Table | | Score ▼ |

☑ Author retains unrestricted copyrights and publishing rights

Average publication time: 52 weeks

Maximum publication fees (APCs): 1000 €

Select subject ❓

| All ▼ |

Select language

| English ▼ |

Select keywords ❓

| Search keywords | ✕ |

☐ academic libraries
☐ advertising discourse
☐ advocacy
☐ allergology
☐ amputation
☐ archaeology
☐ archival science
☐ archival studies
☐ archive studies
☐ arrhythmia
☐ art
☐ artificial nutrition

✖ Reset filters

According to your input, we automatically set the language filter to "English". You can change it along with the other filters. Setting the subject filter to "Bibliography. Library science. Information resources" will likely result in more specific results. ✕

Score: 100% ▼

Quantitative Science Studies

Publisher: The MIT Press

Keywords: scholarly communication, science indicators, science policy, scientific workforce, system of science

Quantitative Science Studies

Publisher:	The MIT Press
ISSNs:	2641-3337
Average publication time:	22 weeks
Language(s):	English
Author retains unrestricted rights:	✓
Article receives DOI:	✓
Peer review method:	Blind peer review
License(s):	CC BY
Maximum publication fees (APCs):	800 USD (approx. 761.9€)
Plan S compliance:	✓
In DOAJ since:	7/21/2020

Website

Author instructions

Aims & scope

Editorial Board

Subjects:

Science (General)

Keywords:

scholarly communication   science indicators

science policy   scientific workforce

system of science

Give us feedback

Resonanz zur Betaversion

- Funktionalität
  - Empfehlungsalgorithmus
  - Score-Hintergrundinformationen
  - zugrundeliegende Daten

- Usability/User Interface
  - Filter
  - Ergebnisliste
  - Eingabemaske
  - Hilfetexte/Erläuterungen
  - Journal-Profilseite

- (Design)

Photo: PD-USGov-Interior-FWS, Public domain, via Wikimedia Commons

4. Im Inneren des B!SON
– technische Verfahren hinter den Empfehlungen

BISON: Struktur

- Front-end

- Back-end

- Deployment

- Extension
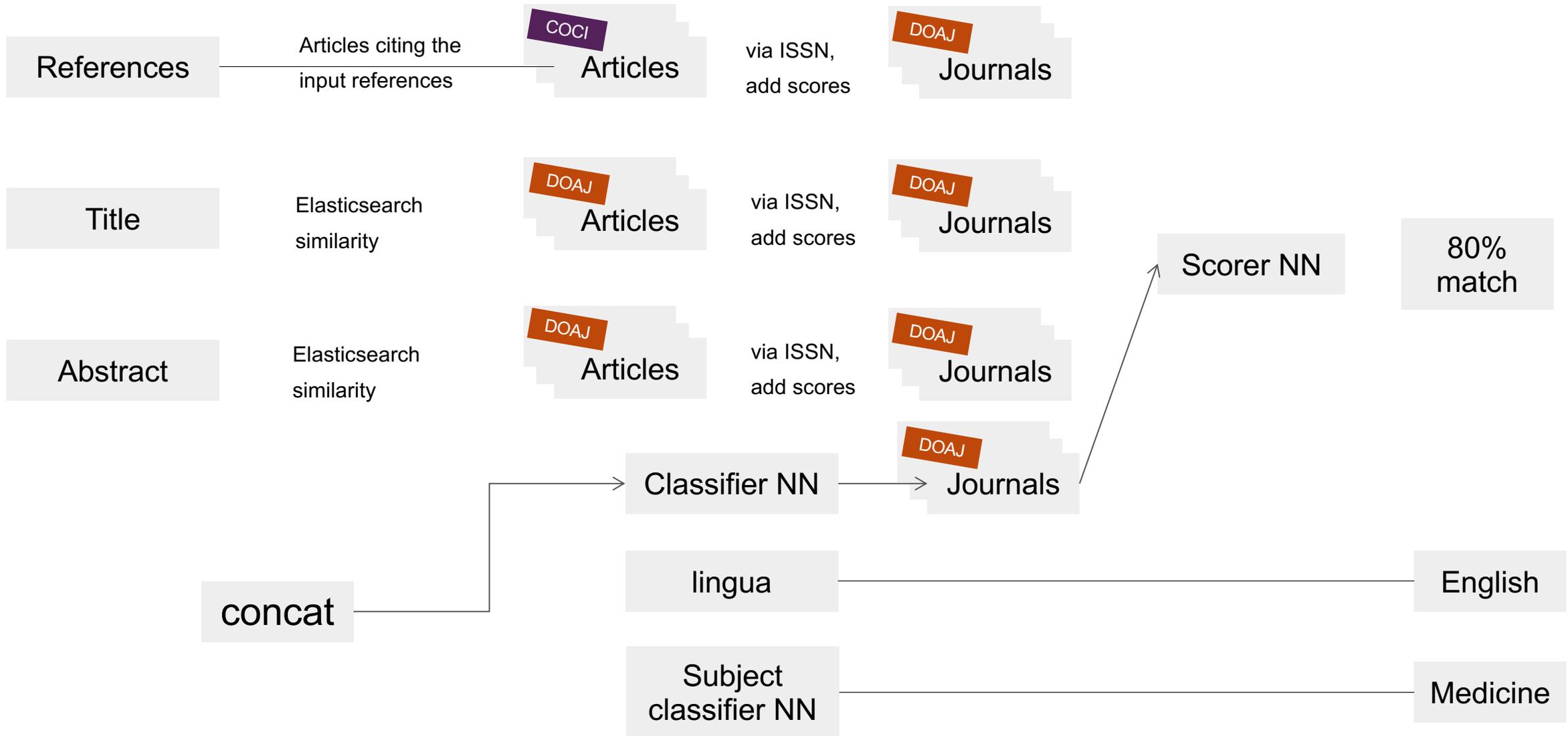
EIN PROJEKT VON TIB & SLUB

Front-end

- Vue.js-Projekt

- Verfügbar unter https://gitlab.com/TIBHannover/bison/buffalo

Back-end

- Django-Application (Python)

- Verfügbar unter https://gitlab.com/TIBHannover/bison/yak

- Elasticsearch und PostgreSQL als Datenbanken

- DOAJ und COCI (von OpenCitations) werden komplett heruntergeladen und eingelesen

- Plan-S compliance über Journal Checker Tool API abgefragt

- Aktuelle Anforderungen:

  - 6 GB GPU Arbeitsspeicher

  - 500GB Festplattenspeicher

  - 25GB Arbeitsspeicher (stark von Konfiguration abhängig)

Recommendations

| References | Articles citing the input references | COCI Articles | via ISSN, add scores | DOAJ Journals |

| Title | Elasticsearch similarity | DOAJ Articles | via ISSN, add scores | DOAJ Journals |

| Abstract | Elasticsearch similarity | DOAJ Articles | via ISSN, add scores | DOAJ Journals |

Scorer NN → 80% match

concat → Classifier NN → DOAJ Journals

lingua → English

Subject classifier NN → Medicine

Deployment

- Ansible ("infrastructure-as-code")

- derzeit nicht Open Source

- Ray für ML Modelle

- läuft auf eigenem Server der TIB

B!SON

5. Lokale Einbindung und Anpassung

Ziel im AP Lokale Implementierung:
B!SON möglichst nahtlos in die bestehenden Beratungsworkflows und Webseiten interessierter Bibliotheken und Forschungseinrichtungen zu integrieren

Zielgruppe:
wissenschaftliche Einrichtungen → Anlaufstellen für Autor:innen
- Bibliotheken
- Fachinformationsdienste

Mehrwert:
- Nutzende können den vollen Funktionsumfang von B!SON innerhalb der lokalen Webservices nutzen
- Keine Entwickler-/API-Kenntnisse nötig
- Recommender individuell konfigurierbar (lokale Policies)

Umsetzung

1. Entwicklung Prototyp TYPO3-Extension
   - Voraussetzung an den Einrichtungen: Kenntnisse für die Einrichtung von Typo3-Extensions (keine Softwareentwicklung/API-Kenntnisse nötig)

2. Ermöglichung einer Anpassung der Empfehlungen an lokale Policies oder einrichtungsspezifische Präferenzen, z.B.:
   - Setzen von Kostenobergrenzen
   - Lizenzanforderungen (CC-BY)
   - Auswahl bestimmter Verlage und/oder Journals (Berücksichtigung oder Ausschluss)

B!SON

6. Ergebnisse und Ausblick

Ergebnisse

- Erkenntnisse zu Anforderungen an einen Journal Recommender
- Algorithmus:
  - Bestimmung von zu Eingabedaten ähnlichen Artikeln / Journals
  - Erkenntnisse zur Eignung verschiedener semantischer & bibliometrischer Verfahren

- **Zentraler, unabhängig nutzbarer Dienst „B!SON Journal Recommender"**
- **Extension für lokale Einbindung (Prototyp)**

- Netzwerk das Projekt unterstützender Einrichtungen
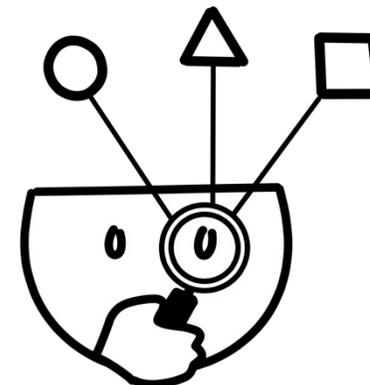
Ausblick

- Herausforderungen:
  - Datenqualität/-aktualität/-vollständigkeit
  - Algorithmus-Finetuning
    - Schwellwerte für bibliometrische und semantische Verfahren
    - Normalisierung der bibliometrischen Verfahren
    - Gewichtung der Ähnlichkeitsmaße für Titel/Abstract/Referenzen
  - Bekanntheit bei Wissenschaftler:innen

- B!SON als Basistechnologie
  - weitere Anwendungsfälle
  - andere Datenquellen
- B!SON an den Einrichtungen

B!SON

Vielen Dank!

B!SON Journal Recommender: https://service.tib.eu/bison

Projektwebsite: https://projects.tib.eu/bison

Kontakt: bison@tib.eu

Icons: https://doi.org/10.5281/zenodo.5608845, Creative Commons Zero https://creativecommons.org/publicdomain/zero/1.0/

EIN PROJEKT VON

TIB

SLUB

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung

FÖRDERKENNZEICHEN: 16TOA034A

Projektpartner

TIB

SLUB
Wir führen Wissen.

Datengebende Partner

DOAJ

OpenCitations

Förderung

Bundesministerium
für Bildung
und Forschung

Förderlinie: „Beschleunigung der
Open-Access-Transformation"

Unterstützende Einrichtungen