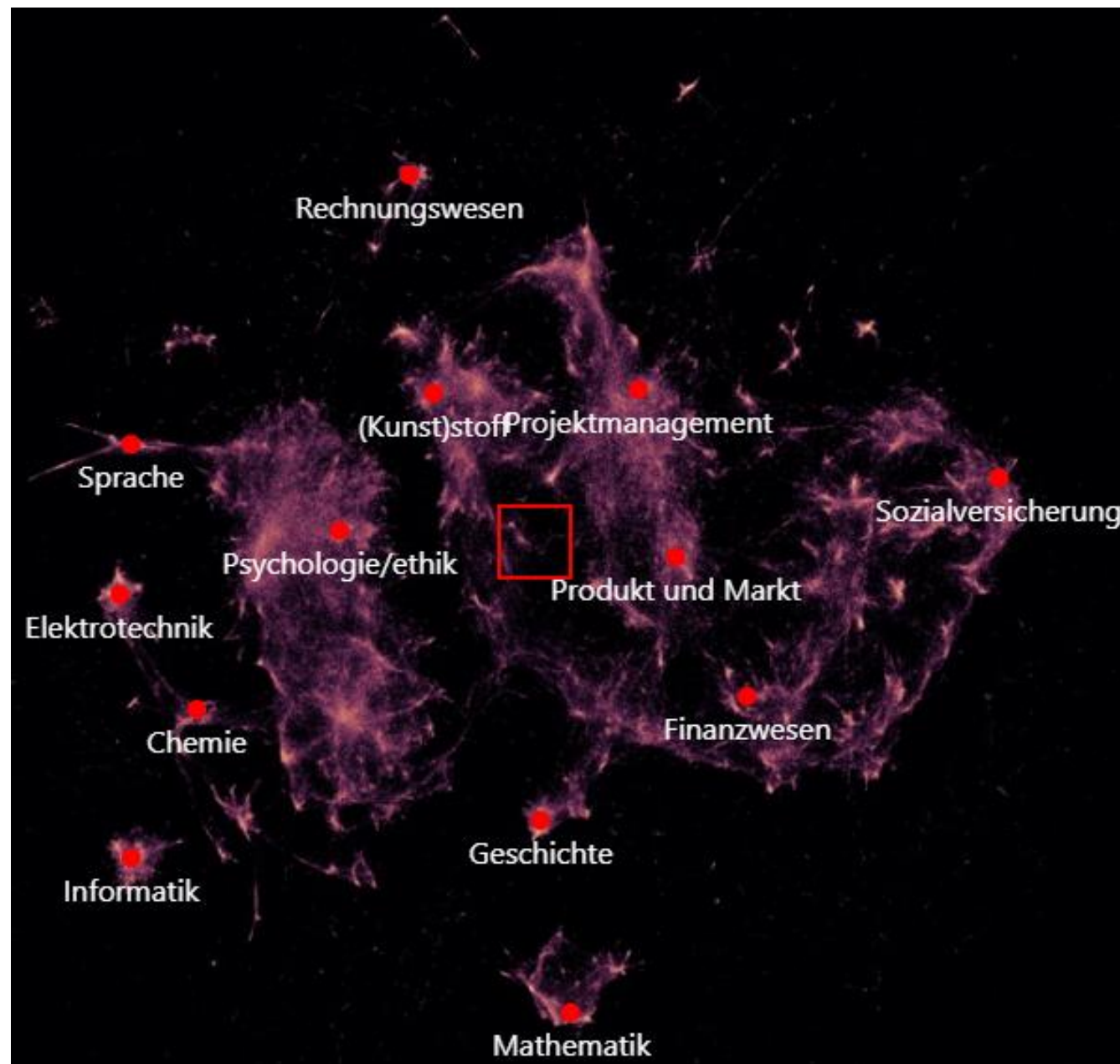




machine learning
excellence delivered

AI-based search and navigation for book collections



Outline



- Motivation
- Implementation principle
 - Word map: keywords and topics
 - Document map
 - Search clarification
- Search queries
- Applications

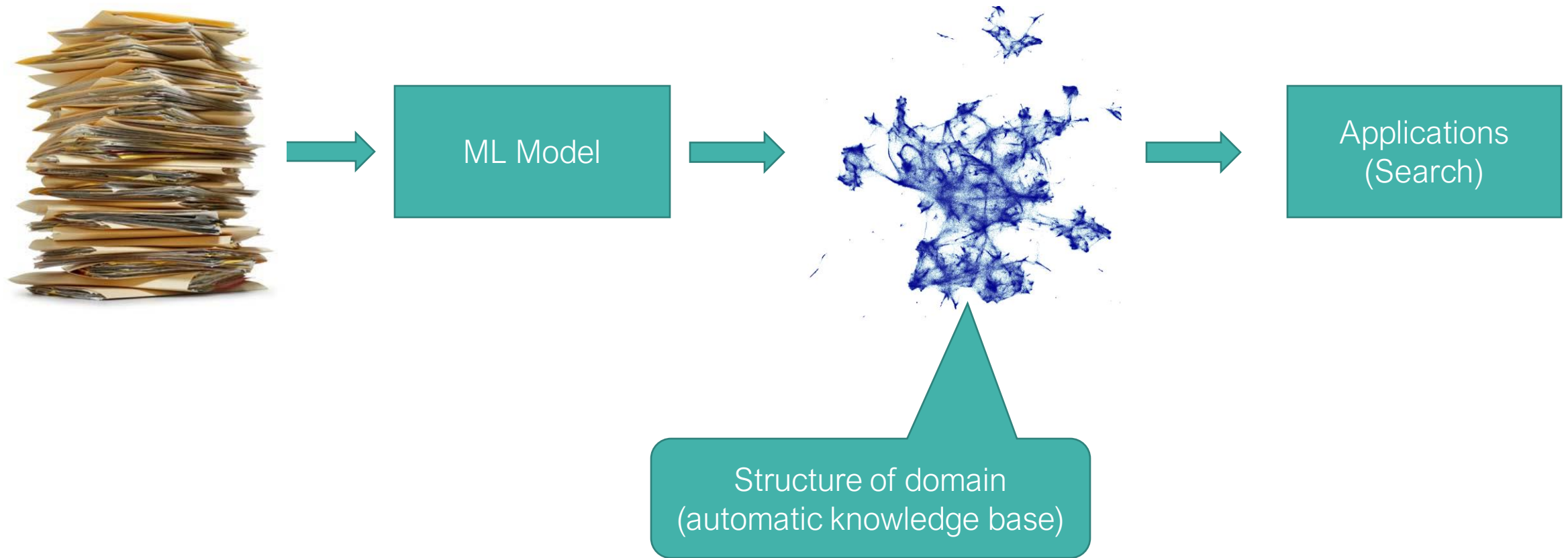
Motivation



- Knowledge workers spent up to **2.5 hour per day*** to find the information
 - Equals to **€25 million per year** for 1000-workers corporation
 - Extra cost of **not finding** important data
- Issues:
 - Varying/special vocabulary
 - Duplicate documents
 - Large number of results
 - Lack of user data ('clickstream') in corporate and special search systems

* Source: <http://intranetfocus.com/2-5-hours-a-day-spent-searching-really/>

Implementation: semantic map

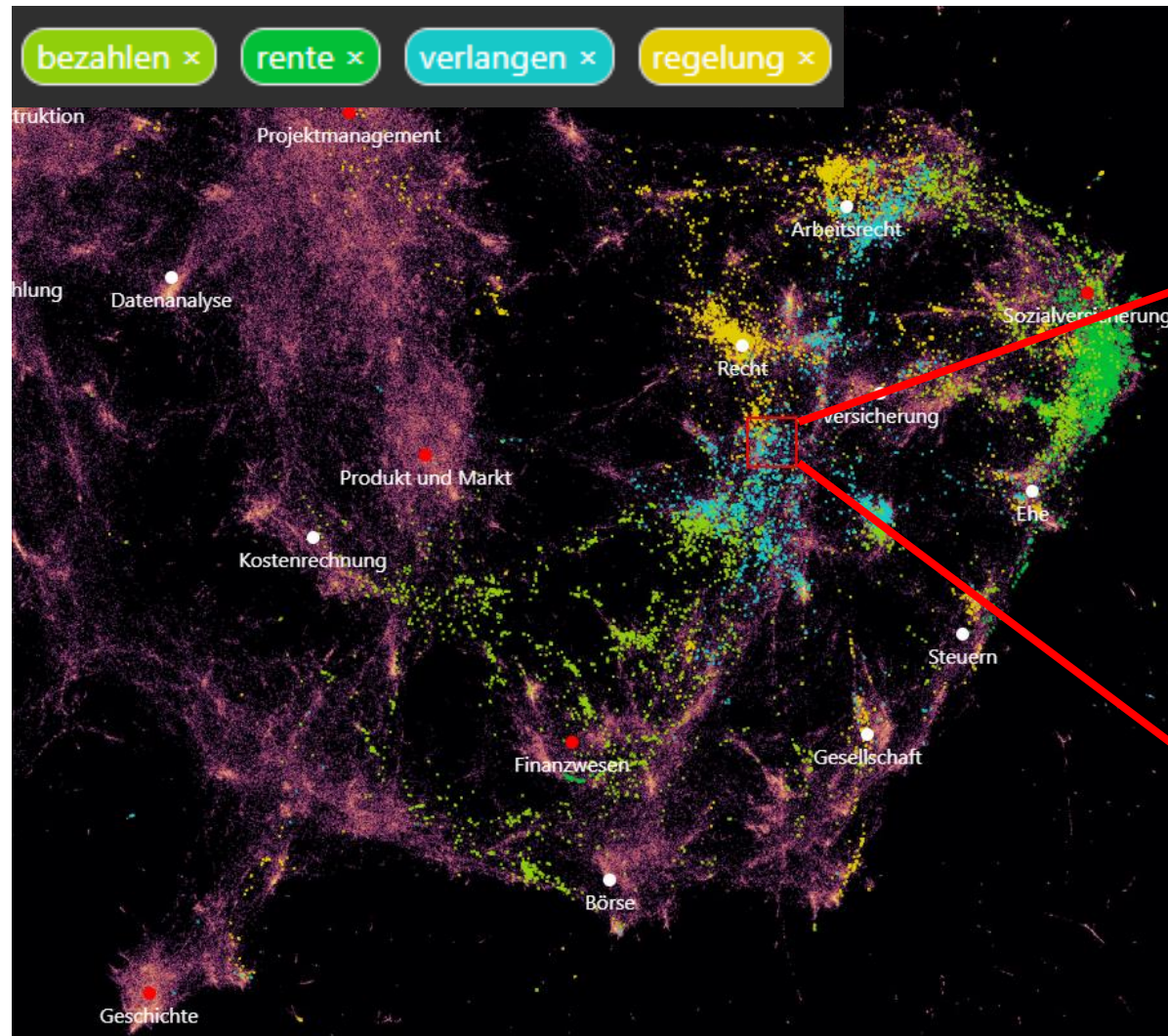


Data for demos



- Books from Edubase platform
 - <https://app.edubase.ch/>
 - 2694 German-language books
 - 430217 pages in total
 - Search demo: <https://edubase.silkdata.ai/>
- Scientific papers related to COVID-19 crisis
 - Source of papers: <https://www.semanticscholar.org/cord19>
 - Free access
 - 53000 papers in our demo
 - Search demo: <https://covid19.silkdata.ai/>

Word map



ARTICLES ①

Allgemeine Vertragslehre 1/2

<https://app.edubase.ch/#doc/55442>

Allgemeine Vertragslehre 1/2

<https://app.edubase.ch/#doc/37642>

Recht – TK 2019

<https://app.edubase.ch/#doc/21697>

Tatbestand Recht - Band 1

<https://app.edubase.ch/#doc/22937>

Tatbestand Recht Bd. 1

<https://app.edubase.ch/#doc/57017>

ICT-Rechtskonformität

<https://app.edubase.ch/#doc/13965>

Rechtskunde

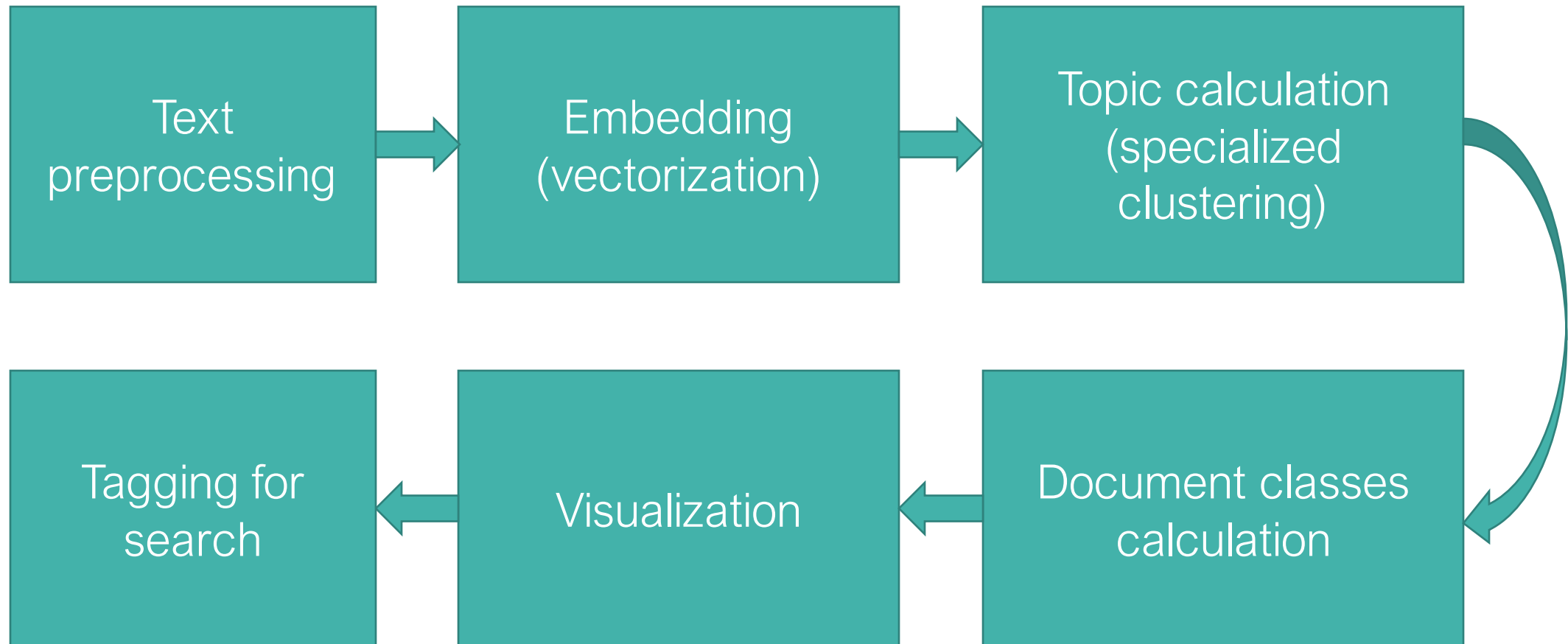
<https://app.edubase.ch/#doc/27605>

Ergänzung des Schweizerischen Zivilgesetzbuches

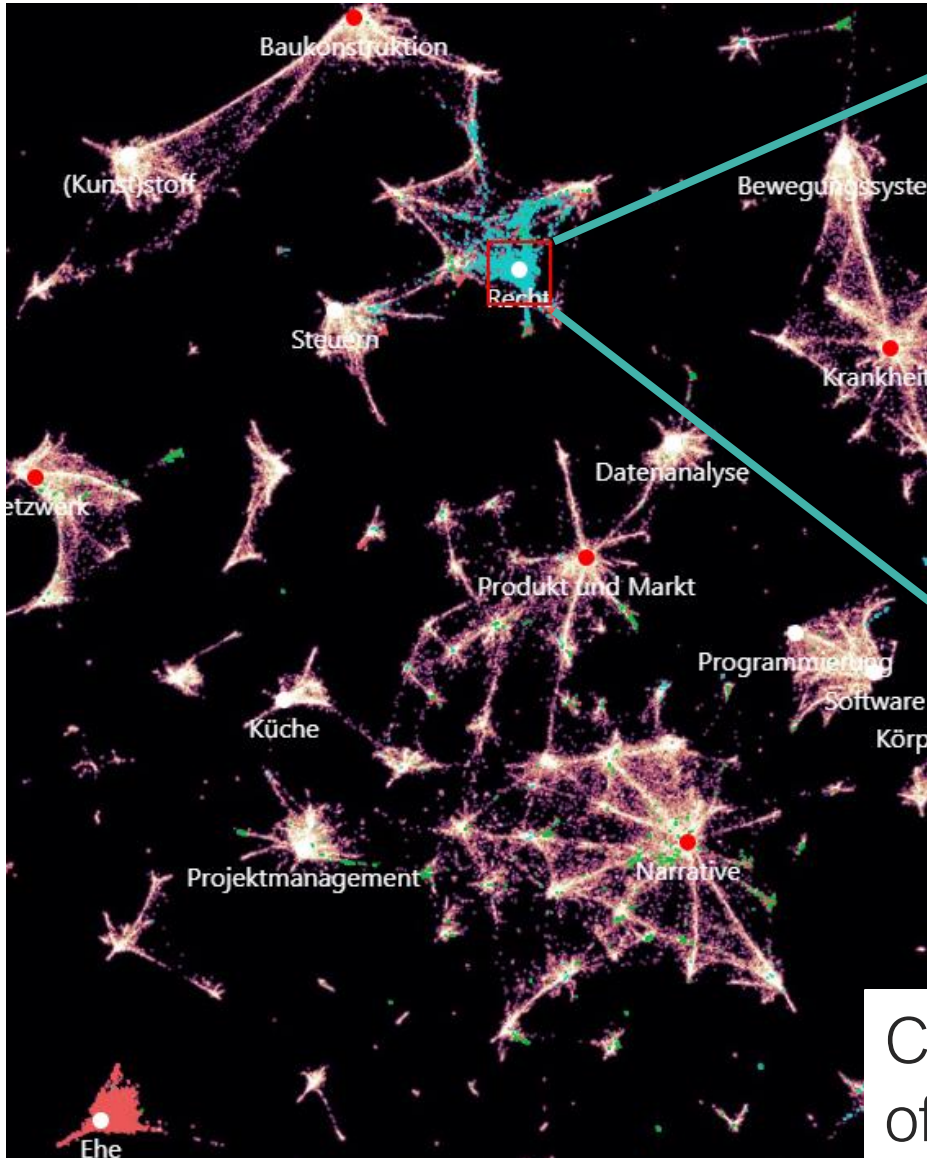
<https://app.edubase.ch/#doc/19897>

Connection between keywords, topics and documents (snippets)

NLP pipeline



Document map



ARTICLES ⓘ

Internationales Zivilprozessrecht, p.465
<https://app.edubase.ch/#doc/35910/465>

Schweizerisches Obligationenrecht - Allgemeiner Teil, p.323
<https://app.edubase.ch/#doc/31654/323>

Internationales Zivilprozessrecht, p.352
<https://app.edubase.ch/#doc/35910/352>

Allgemeines Verwaltungsrecht, p.271
<https://app.edubase.ch/#doc/35914/271>

Schweizerisches Obligationenrecht. Allgemeiner Teil I & II, p.323
<https://app.edubase.ch/#doc/31974/323>

Nachbarschaft - was gilt im Konfliktfall, p.44
<https://app.edubase.ch/#doc/40266/44>

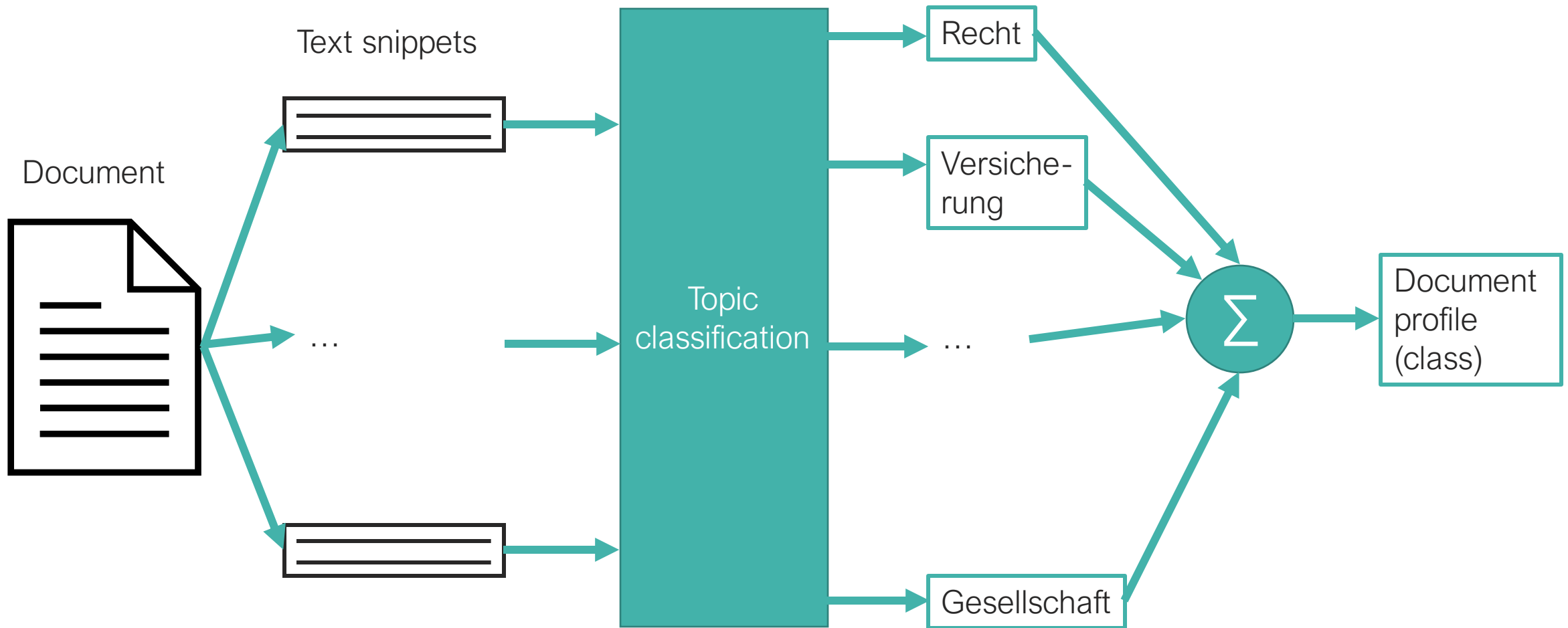
ZGB. Das Schweizerische Zivilgesetzbuch, p.748
<https://app.edubase.ch/#doc/31986/748>

ZGB. Das Schweizerische Zivilgesetzbuch, p.101
<https://app.edubase.ch/#doc/31986/101>

Allgemeines Verwaltungsrecht, p.615
<https://app.edubase.ch/#doc/35914/615>

Connection between document classes (groups of topics) and documents (book pages)

Document map: principle



Search

Search bar: Search

<https://app.edubase.ch/#doc/50093/46>
Rückkühlung in der Kältetechnik, p.46
Allgemein In diesem Kapitel wird lediglich die **Regelung** zur Einhaltung der Rückkühlbedingungen,... nicht aber die **Regelung** der Gesamtanlage behandelt. Mit der Leistungsregelung der Rückkühler bezweckt

Heizung/kühlung Find similar

<https://app.edubase.ch/#doc/3257/63>
Leben in der Schweiz, p.63
Regelung betreffend Rauchen in unserem Kanton? A Gibt es Unterschiede zu den Nachbarkantonen? A... Seit wann gilt diese **Regelung**? A Wie war es vorher? A Erwarten Sie in den nächsten Jahren Änderungen?

Steuern Find similar

<https://app.edubase.ch/#doc/37234/63>
Modul 5: Leben in der Schweiz, p.63
Regelung betreffend Rauchen in unserem Kanton? Gibt es Unterschiede zu den Nachbarkantonen? Seit wann... gilt diese **Regelung**? Wie war es vorher? Erwarten Sie in den nächsten Jahren Änderungen? Wenn ja: In

Steuern Find similar

<https://app.edubase.ch/#doc/14765/142>
IT Security Management, p.142
Regelungen für Wartungs- und Reparaturarbeiten Verhinderung ungesicherter Netzzugänge **Regelung** zum Umgang mit Betriebsmitteln **Regelung** der Verantwortlichkeiten, Aufgaben

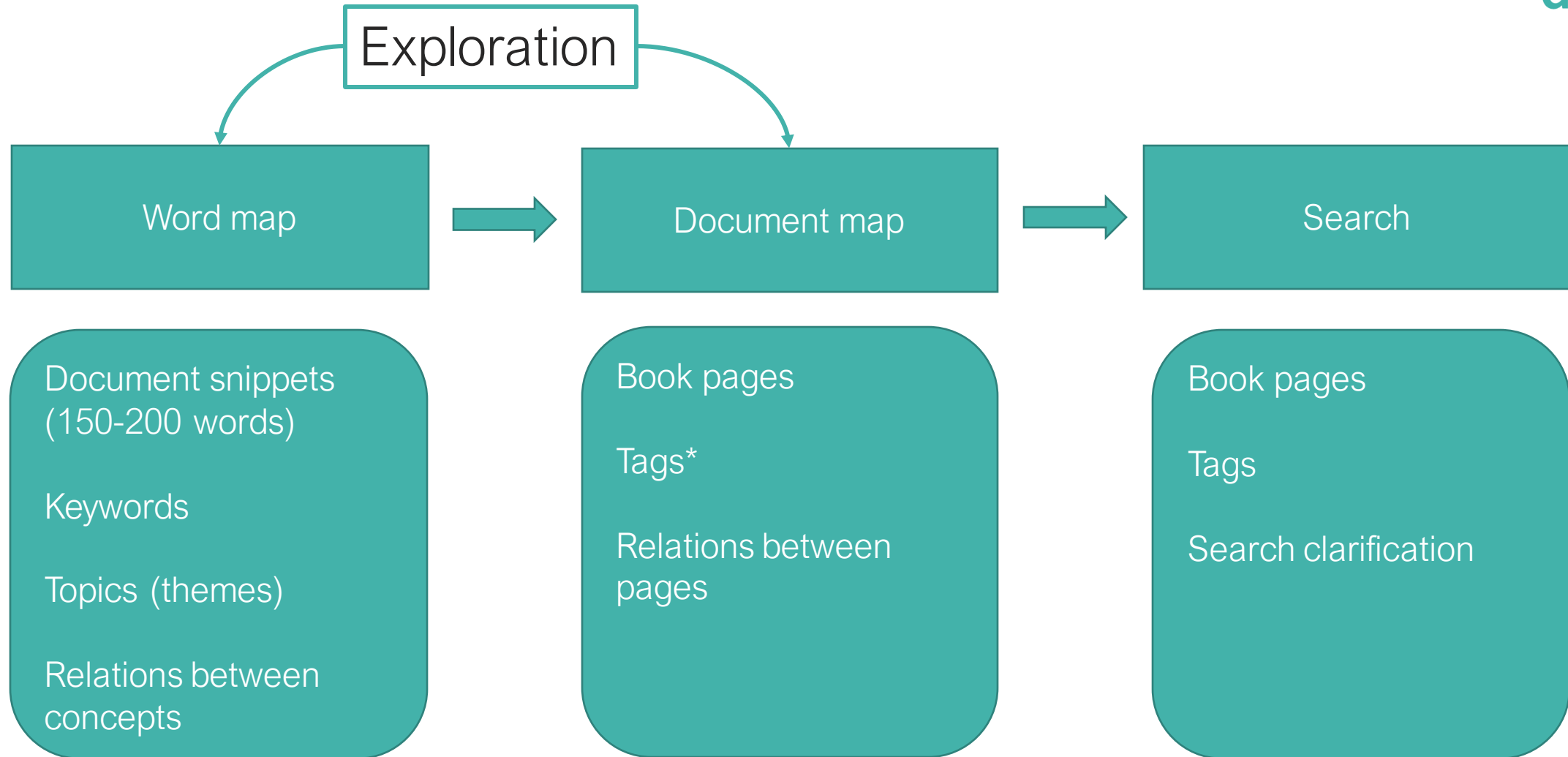
Service management

'Find similar':
looks for semantically similar
documents (in topic space)

Document classes (contexts) for search clarification

Recht	Arbeitsrecht	Ehe	Gesellschaft	Versicherung	Sozialversicherung	Elektrotechnik	Heizung/kühlung	Steuern	Kauf
Eigentumsrecht	Strafrecht	Projektmanagement	Gericht	Staat	Schuldner und gläubiger	Persönlichkeitschutz	Wohnung/Miete		
Netzwerk	Bildung	Team	Geschichte	Rechnungslegung	Baunormen	Recycling/reinigung	Baukonstruktion	Kredit/Hypothek	

Principle at a glance



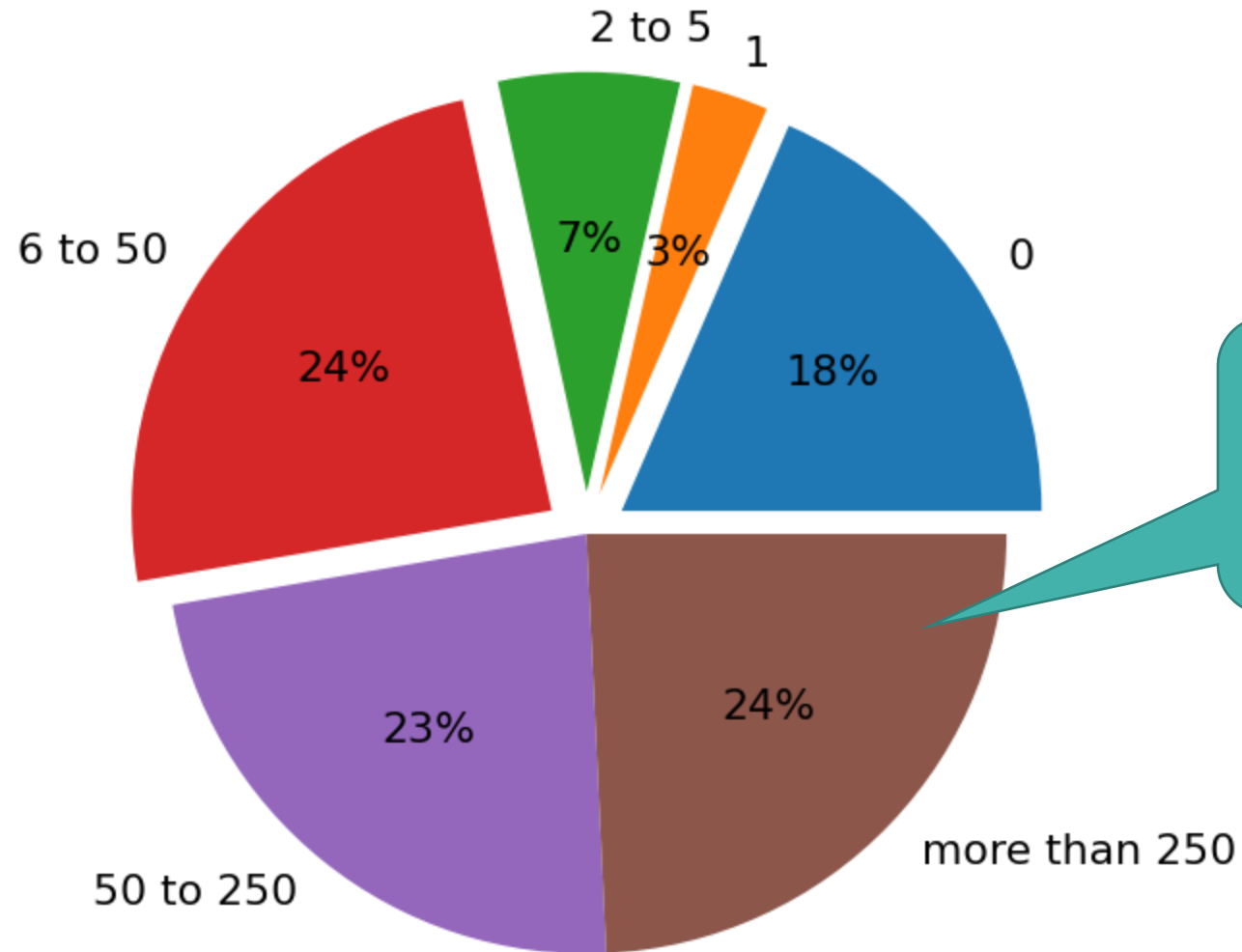
* Document classes, combinations of topics; **single tag** per document (book page)

Search queries



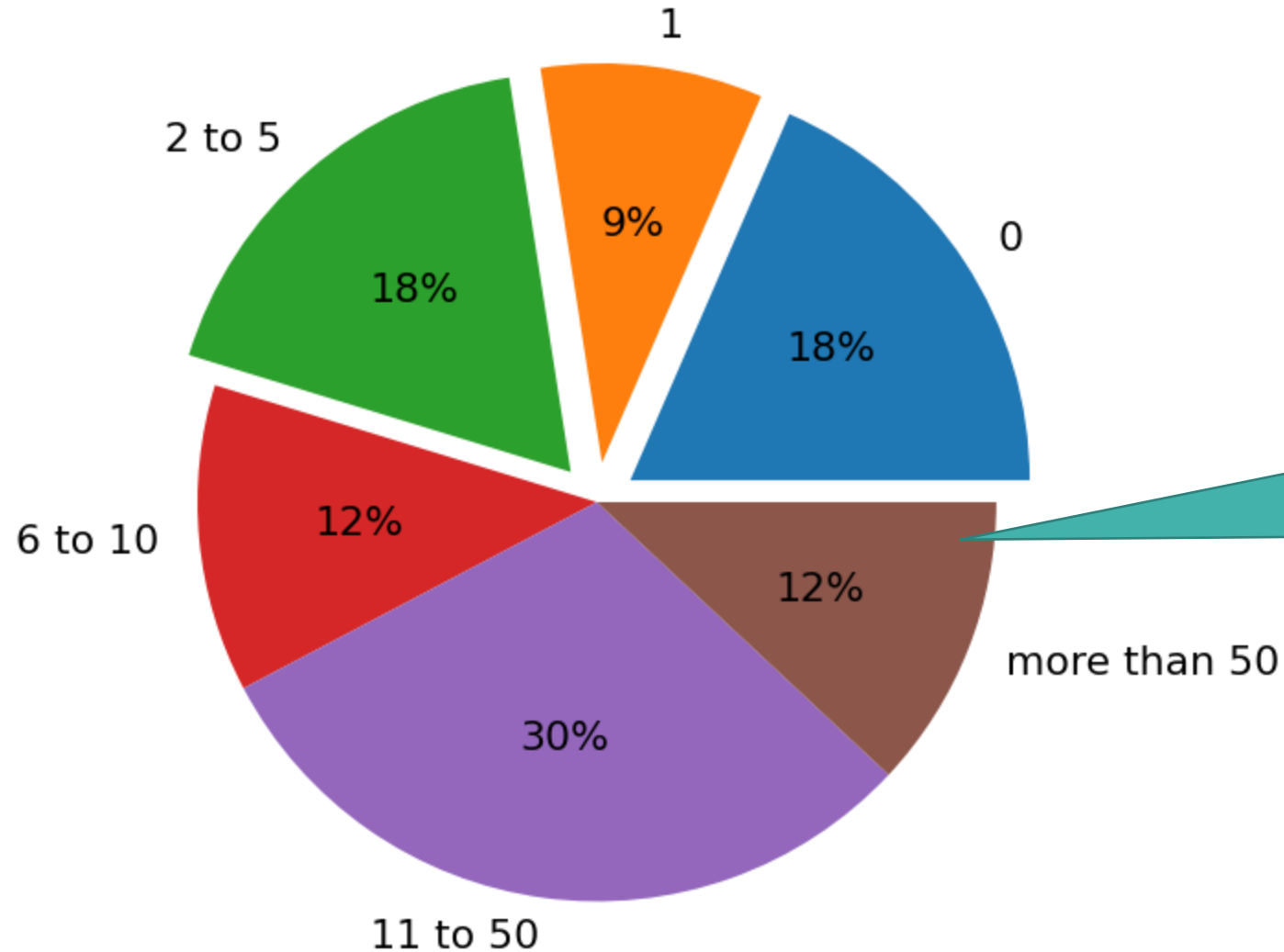
- Edubase platform
- Collected from Nov 2019 till Aug 2020
- 310537 queries in total
- 1650 per day
- 58889 unique queries
- 1 to 31 words (99% - 8 words or less)
- 2 to 234 symbols (99% - up to 63 symbols)
- Up to 300,000 results per query
(25% queries yields more than 230 results)
- Up to 104 document classes

Number of results per query



Almost half of (unique) queries have more than 50 results

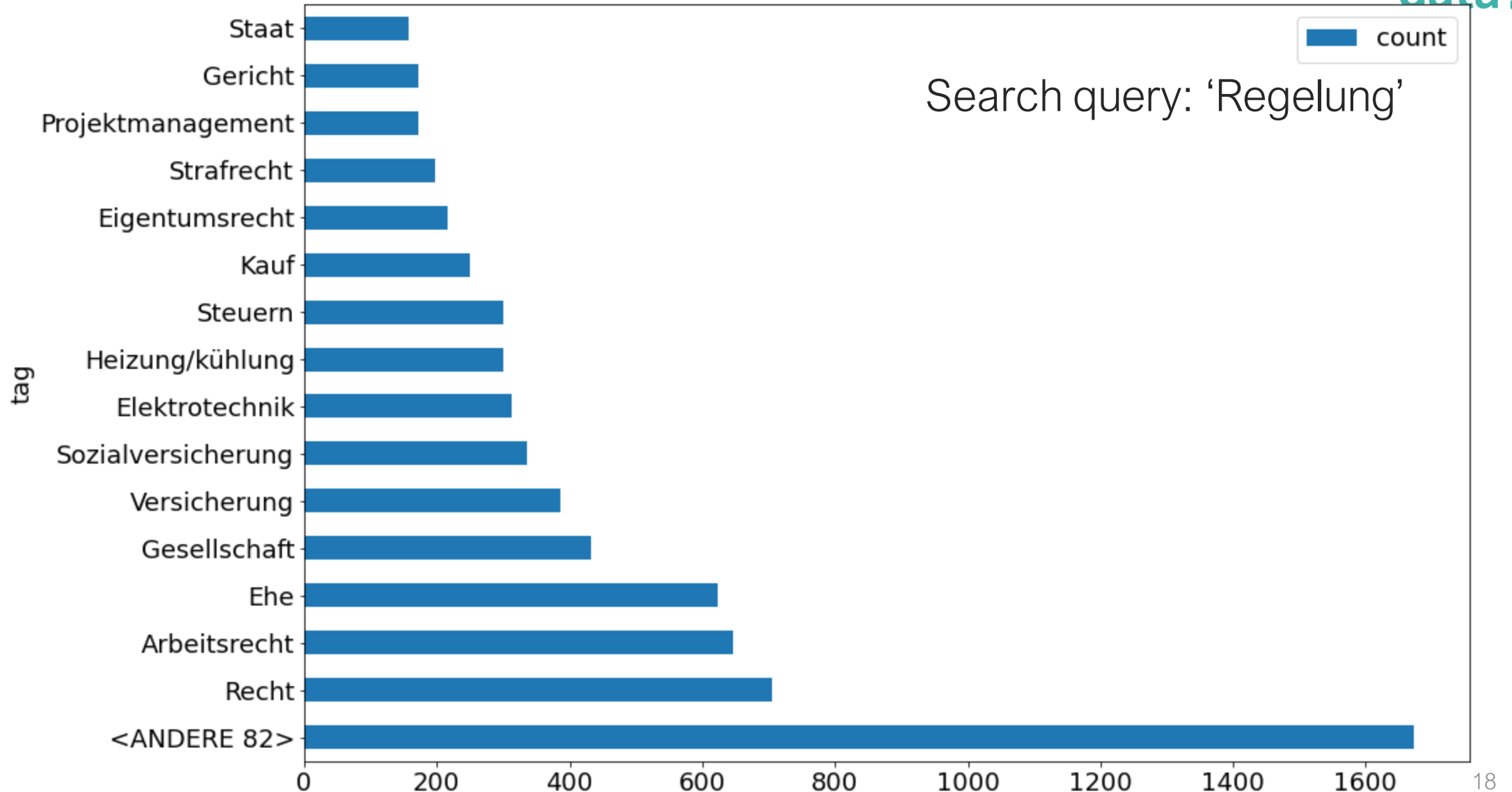
Number of tags per query



Usual solution

Add more keywords to
search query

Tag histogram



Our proposal

Provide automatic
suggestions to narrow
down search

Search clarification

Full: 688 pages
Recht: 71 pages
Ehe: 63 pages



Search bar: Search

<https://app.edubase.ch/#doc/50093/46>
Rückkühlung in der Kältetechnik, p.46
Allgemein In diesem Kapitel wird lediglich die **Regelung** zur Einhaltung der Rückkühlbedingungen,... nicht aber die **Regelung** der Gesamtanlage behandelt. Mit der Leistungsregelung der Rückkühler bezweckt
Heizung/kühlung Find similar

<https://app.edubase.ch/#doc/3257/63>
Leben in der Schweiz, p.63
Regelung betreffend Rauchen in unserem Kanton? A Gibt es Unterschiede zu den Nachbarkantonen? A... Seit wann gilt diese **Regelung**? A Wie war es vorher? A Erwarten Sie in den nächsten Jahren Änderungen?
Steuern Find similar

<https://app.edubase.ch/#doc/37234/63>
Modul 5: Leben in der Schweiz, p.63
Regelung betreffend Rauchen in unserem Kanton? Gibt es Unterschiede zu den Nachbarkantonen? Seit wann... gilt diese **Regelung**? Wie war es vorher? Erwarten Sie in den nächsten Jahren Änderungen? Wenn ja: In
Steuern Find similar

<https://app.edubase.ch/#doc/14765/142>
IT Security Management, p.142
Regelungen für Wartungs - und Reparaturarbeiten Verhinderung ungesicherter Netzzugänge **Regelung** des... Passwortgebrauchs **Regelung** zum Umgang mit Betriebsmitteln **Regelung** der Verantwortlichkeiten, Aufgaben
Service management Find similar

Search bar: Search

Recht ×

<https://app.edubase.ch/#doc/35550/59>
Europarecht II, p.59
Warenverkehrsfreiheit burgischen **Regelung** mit den Artikeln 28 und 30 EGV (jetzt Art. 34 und 36... ist, ob eine **Regelung** der streitigen Art den innergemeinschaftlichen Handelsverkehr unmittelbar oder
Recht Find similar

<https://app.edubase.ch/#doc/35550/115>
Europarecht II, p.115
von Dienstleistungen beeinträchtigenkann. (10) Dazu trägt der SDX vor, dass die betreffende **Regelung**... ist die **Regelung** auf sie anwendbar. (13) Wenn selbständige Fremdenführer aus einem Mitgliedstaat im
Recht Find similar

<https://app.edubase.ch/#doc/35966/22>
Staatsrecht der Schweizerischen Eidgenossenschaft, p.22
ausdrücklichen **Regelung** des Bundes 325 3. / VR) P) NQQ m & Q RNIR abschliessende **Regelung** des Bundes... in der Sache 326 4. / VR) P) NQQ auf eine abschliessende **Regelung** des Bundes im Verfahren 327 XX \\
Recht Find similar

Similar documents

Q regelung x Search

Regenerative Energiesysteme, p.313 x

<https://app.edubase.ch/#doc/56832/313>
Regenerative Energiesysteme, p.313
vornehmen, für die man zwischen den zwei Verfahren Stall-**Regelung** und Pitch-**Regelung** unterscheidet. Bei... der Stall-**Regelung** wird ausgenutzt, dass es bei großen Anströmwinden zum Strömungsabriss (stall)

Elektrotechnik Find similar

<https://app.edubase.ch/#doc/41186/294>
Regenerative Energiesysteme, p.294
man zwischen den zwei Verfahren • Stall - **Regelung** und • Pitch - **Regelung** unterscheidet. Bei der Stall... - **Regelung** wird ausgenutzt, dass es bei großen Anstellwinden zum Strömungsabriss (stall) kommt

Luftfahrt Find similar

<https://app.edubase.ch/#doc/54858/648>
Fachkunde Industrielektronik und Informationstechnik, p.648
7. 6. 2 Windkraftanlagen 647 **Regelung** Ist die Windstärke für eine Windkraftanlage zu groß, muss der... gedreht werden. Man unterscheidet da bei die Pitch - **Regelung** und die Stall - **Regelung** (Tabelle 1).

Elektrotechnik Find similar

<https://app.edubase.ch/#doc/34139/28>
Vertiefung Fachkunde Gebäudetechnikplaner/innen EFZ Heizung, p.28
Regelung ist eine Einheit, die eine oder mehrere Regelgrößen und diese in einem Gerät
..... 125. Worin unterscheiden sich eine **Regelung** und eine Steuerung?

Elektrotechnik Find similar

<https://app.edubase.ch/#doc/40538/205>
Elektrische Mess- und Regelungstechnik, p.205
aus. Je nach der Art der **Regelung** und den Qualitätsanforderungen an die **Regelung** treten verschiedene... einer **Regelung** dargestellt. Hier wird auch noch einmal bildlich der Unterschied zwischen einer

Elektrotechnik Find similar

<https://app.edubase.ch/#doc/54858/549>
Fachkunde Industrielektronik und Informationstechnik, p.549
548 5. 4. 6. 1 **Regelung** von P - Strecken 5. 4. 6 Regelkreise (Beispiele) 5. 4. 6. 1 **Regelung** von P... bleibt wirkungslos und das Regelverhalten wird nicht linear. Der PÜD - Regler wird bei **Regelung** auf

Elektrotechnik Find similar

Document vectors is based on topic distribution

Applications



- Search improvement
- Search on specialized data (where Google and Co. cannot be used)
- Exploratory search
- Ontology engineering
- ‘Visual contents’ of a book
- Visual topic model

Possible data for search and navigation:

- Books/journals (publishers)
- Internal document collections (research, corporate, ...)
- Emails
- Legal data (laws, regulations, contracts, standards)

Our demo: <https://edubase.silkdata.ai/>

Thank you for attention!



visit: silkdata.ai

mail to: hello@silkdata.ai



www.silkdata.ai
hello@silkdata.ai
+48 452 380 167

Silk Data Sp. z o.o.

Domaniewska 17/19, off. 133
02-672 Warsaw,
Poland
+48 452 380 167

SilkCode GmbH

Luisenstraße 62
D-47799 Krefeld
Germany
+49 (2151) 387 3531