

KULTurSUchmaschine: Same but different... Information Retrieval auf Museumsdaten

Katharina Faßbauer¹, Lea-Kathrin Gleißner¹, Amina Ismael¹ und Elias
Messner¹

Universität Leipzig, 04109 Leipzig, Deutschland

Zusammenfassung. Information Retrieval ist eng mit der Domäne des Kulturerbes verbunden. Dennoch hat sich die Datengrundlage in jüngerer Zeit von Objektdaten zu Volltexten gewandelt. Museen digitalisieren ihre Bestände und schaffen so durchsuchbare Korpora, an die vielfältige Bedürfnisse gestellt werden. **Kusu** ist ein Vorschlag über Ansätze zur Befriedigung dieser Bedürfnisse nachzudenken. **Kusu** zeichnet sich durch eine farbasierte Erweiterung (Content Based Information Retrieval) zum klassisch textuellen Zugang zu offen lizenzierten Museumsdatenkorpora aus. Es ist mit der *Elasticsearch*-Engine implementiert und arbeitet auf Bilddaten und den speziellen LIDO-XML-Daten der Museen. Es stellt sich heraus, dass die Query-Verarbeitung, im Gegensatz zur Index-Konfiguration, keinen signifikanten Einfluss auf die Retrieval-Qualität hat.

Schlüsselwörter: Cultural Heritage · Information Retrieval · Museum Data Retrieval · OpenGLAM · XML-Retrieval

1 Einleitung

Da sich das Information Retrieval aus den Bibliothekswissenschaften entwickelte und noch heute Problemstellungen mit ihnen teilt, [5,1] ist die fachliche Nähe zu dieser GLAM-Institution¹ eindeutig. Weniger Relevanz für das IR hingegen hat das „M“ in „GLAM“. Dennoch gehen mit der Institution Museum digitalisierte Katalogisierungs-, Klassifizierungs- und Erfassungsprozesse einher, die wiederkehrend diskutiert und kritisiert werden.

Das moderne IR entfernte sich von kategorisierten Daten, [14] denn für die moderne Textanalyse sind die Inhalte von Bibliotheksbüchern interessanter als ihre Metadaten. Museen hingegen generieren seit einiger Zeit *en masse* in Digitalisierungsprojekten kategorisierte Objektdaten, die die Basis für die Suchdomäne „Kulturerbe“ bilden. Präsenz und Dringlichkeit der Digitalisierung im Museumswesen tragen zur Informationsflut bei,² die das IR seit Jahren herausfordert. [9] Objektdaten unterscheidet von Volltexten eine große Heterogenität im Datenformat (strukturierter und unstrukturierter Text, sowie Multimedia in Bild und

¹ GLAM = Galleries, Libraries, Archives, Museums

² vgl. hierzu die Präsenz des Themas [auf der Website des deutschen Museumsbundes](#) und unter dessen Publikationen, [Diskussionen auf der republica 2019](#) und [6]

Ton), die sie für Sprachverarbeitung weniger gut geeignet macht. [9] Neben heterogenen Datenformaten als Basis zeichnen sich Museumssuchmaschinen durch eine heterogene Nutzendengruppe aus. Interface- und Retrieval-Gestaltung befriedigen auf veröffentlichten Museumsdatenbanken derzeit nur das Informationsbedürfnis von Fachleuten. [6]³ Visuelle Zugänge zur Sammlung, die zusätzlich zu Suchmöglichkeiten *Browsing*-Szenarien implementieren setzen dort an.⁴ Suchmaschinen auf Museumsdaten werden derzeit meist direkt mit der entsprechenden „Sammlung Online“ des jeweiligen Hauses zur Verfügung gestellt.⁵ Initiativen wie die *Europeana*, oder *iArt* machen Kulturgut institutionsübergreifend verfügbar und durchsuchbar. Beide Projekte nutzten offen verfügbare Museumsdigitalisate. [13,7] Die *Europeana* will europäisches Kulturerbe an einem Ort vereinen. [7] *iArt* arbeitet mit Deep-Learning-Techniken zur Bildanalyse um Fachpublikum den Werkvergleich zu erleichtern. [13]

Wir besprechen hier die prototypische Implementierung der Kultur-Suchmaschine **Kusu**. **Kusu** durchsucht museumsübergreifend verschiedene Datensätze, die unter offenen Lizenzen veröffentlicht wurden. Neben semi-strukturiertem Text (alle ausgewählten Datenbanken nutzen das Format LIDO-XML für den Datenexport) wird farbbasiert auf Bilddaten gesucht und somit somit die Multimedialität der Domäne bedient. [6,15]

Um kulturelle Sensibilität gegenüber den Objekten bei der Arbeit mit, und Veröffentlichung von, Museumsdaten zu erreichen, braucht es entsprechendes Domänenwissen.⁶ In diesem Zuge weiterhin relevant ist die Übersetzungslücke zwischen realem Objekt und Digitalisat. [8,18] Da **Kusu** nicht veröffentlicht wird, haben wir aus Kapazitätsgründen die Daten nicht auf diskriminierende und verletzende Inhalte geprüft.

2 Daten

Die für Kusu genutzten Daten stammen verschieden anteilig vom **Museum für Kunst und Gewerbe Hamburg (MKG)**, von der Institution **kult Westmünsterland, Lampensammlung Touché** und dem **Deutschen Museum** (siehe Tabelle 1). Alle verwendeten Datensätze wurden spätestens für **Coding da Vinci (CdV)** offen lizenziert und gehören in das Feld der OpenGLAM. [10] Das genutzte Metadatenformat *LIDO-XML* („Lightweight Information Describ-

³ Das ist besonders dann schade, wenn dadurch Ziele der Digitalisierung der Museen nicht aufgehen. Ziel des MKG ist bspw. durch Digitalisierung zur Nachnutzung der Objektdaten anzuregen. [11] Nachnutzende Kunstschaffende bemängeln hingegen u.a. die Durchsuchbarkeit von digitalisierten Sammlungen im Allgemeinen. [15]

⁴ Betrachtenswert dazu sind [Publikationen](#) und [Arbeiten](#) um Marian Dörk.

⁵ Beispiele auf den von uns genutzten Daten sind die [Implementierung des Museums für Kunst und Gewerbe](#), die [Implementierung des Deutschen Porträtarchivs](#) und der [genutzte Ableger von museumdigital für kult.westmünsterland](#)

⁶ siehe dazu zum Beispiel die [ethischen Leitlinien der Online Collection](#) der Staatlichen Kunstsammlungen Dresden.

ing Objects“) wurde für museale Bedürfnisse entworfen.⁷

Als internationaler Nachfolger des deutschen Schemas *museumdat* ist LIDO-XML der Standard für die Bereitstellung von Museumsdaten [4]. Trotz der Standardisierung sind im LIDO-XML wenige Felder obligatorisch [4], sodass die verwendeten Daten heterogen bleiben und Parser vor Probleme stellen. Im Allgemeinen löst *LIDO-XML* nicht die vorhandenen Interoperabilitätsprobleme bei Daten unterschiedlicher Museen. [9] Die verwendeten Daten enthalten größtenteils stichwortartige Einträge in den entsprechenden LIDO-Kategorien (siehe Beispiel B).

Der von uns entwickelte Parser wandelt eine XML Datei in ein beliebig tief geschachteltes Python Dictionary um. In diesem sucht er dann für gegebene Key-Listen die entscheidenden Felder auf. Wir beschränken uns beim Indexen auf Felder, die Wörter, Phrasen, Volltexte oder auch Jahreszahlen beinhalten, nämlich Titel, Klassifikation, Work-Type, Beschreibung, Maße sowie Ereignisse und damit verbundene Jahreszahlen oder Personen. Wir extrahieren und speichern auch nicht-Text Felder, wie URL oder IMG-URL. Diese dienen nur für die Präsentation der Suchergebnisse, nicht für die Index-Terme.

3 Suchmaschinen-Design

Nutzer*innen stellen verschiedenste Ansprüche an Museumsdaten, von denen **Kusu** einige befriedigen möchte. Die dafür genutzten Informationsbedürfnisse sind *intellektuelle Gründe*, *Inspirationssuche* und *Besuchsplanung*. Sie orientieren sich an Untersuchungen des MET Museums zum Nutzungsverhalten der MET Online Collection. [16,17] Da große Teile der Nutzer*innen direkt durch die Websuche (mit Google) auf die Einzelobjektseiten der Online Collection stoßen, kommen nur etwa 15% der digitalen Besucher*innen als Benutzer*innen der zur Verfügung gestellten Suchmaschine in Frage.[16]

Von verbesserten Suchfunktionalitäten profitiert vor allem die intellektuell motivierte Gruppe.[16,17] Inspirationssuchende profitieren unter anderem von visuellen Anreizen, Filteroptionen und einer hoch spezialisierten Suche in der Onlinesammlung. [16] Auf Basis dieser Daten nutzt **Kusu** den in Abb. 1 skizzierten Technology-Stack zur bestmöglichen Befriedigung der Nutzendenbedürfnisse.

ElasticSearch Die **Elasticsearch**-Engine bildet **Kusus** Grundlage. Basierend auf *Apache Lucene*, einer Bibliothek für Volltextsuchmaschinen, funktioniert *ElasticSearch* ebenfalls in weiten Anwendungsfällen auf Volltextdaten [3]. *iArt* verwendet *ElasticSearch*. Die Europeana, sowie die Suchmaschine des MKG nutzen das ebenfalls auf *Apache Lucene* basierende *Apache Solr*. [13,7,12] Somit sind die Systeme in der Domäne des Kulturerbes ertestet.

Die Kommunikation zwischen Python und Elasticsearch erfolgt über eine Request API. Kusu verwendet das Python Package *Elasticsearch-py*, den offiziellen low-level Client für Elasticsearch.

⁷ Eine detaillierte Aufführung der Umsetzung des Datenformats beim Deutschen Museum findet sich [hier](#).

Indexing und Textsuche Elasticsearch erlaubt Daten im JSON Format für das Indexing. Python Dictionaries haben die selbe Struktur wie JSON Dateien. Deshalb gibt unser Lido Parser die geparsten Daten als Python Dictionary zurück. Die Keys sind dabei die Namen der Felder, zu denen ein Wert oder eine Liste von Werten des gleichen Typs vorliegt. Ein weiteres Python Dictionary definiert die **Index-Konfiguration**. Darin bestimmen wir für jedes Feld, ob es sich um ein Text-feld handelt, und ob es für das Retrieval berücksichtigt werden soll. Außerdem wählen wir in der Konfiguration den Analyzer (Stemmer) sowie Ähnlichkeitsmaße des Index.

Bei der Textsuche legen wir den **Query-Mode** fest. Wir unterscheiden dabei zwischen disjunktiven Multiterm-Queries, konjunktiven Multiterm-Queries sowie Phrase-Match.

Eine Kombination aus Index-Konfiguration und Query-Mode ist ein **Variant** unserer Suchmaschine.

Browsing - Zugang über Farben und „Kategorien“ Unter Inspirationssuchenden konnte das MET ein nicht-zielgerichtetes „Browsing“ und ein hohes Interesse am visuellem Zugang zu den Suchergebnissen feststellen. [16] Deutsche Forschung stellt ähnliche Bedürfnisse nach zielloser Interaktion mit Online-Sammlungen fest. [6,2] Künstler*innen fragen die Verantwortlichen direkt „Can people browse and find content?“. [15]

Wir schlagen als einfache Zugänge zur Sammlung die Auswahlmöglichkeiten von einer von 12 vordefinierten „Kategorien“ und ein *Content-Based Image-Retrieval* (CBIR) mithilfe von 17 vorgeschlagenen Farben vor. Die Europeana und *iArt*, sowie die Suchmaschinen des MKG, DigiPortas und museum digitals arbeiten mit verschieden ausgeprägten facettierten Suchen und Suchvorschlägen.⁸ [13,7,10] Die Kategorien **Kusus** ergeben sich aus manuell zusammengefassten Einträgen aus dem LIDO-XML-Tag `category`.

Der farbbasierte Ansatz **Kusus** ist in der Domäne einzigartig. Indexiert wurden pro Objekt⁹ neben den XML-Daten textuelle Farbhistogramme. Diese wurden mit `extcolors` erstellt, einem Python Package, das die RGB-Werte von Bildern pixelweise analysiert. Diese RGB-Werte wurden auf deutsche Farbnamen gemappt. Leider wurden „metallische“ Farben häufig falsch zugeordnet. Für jedes Bild wurde eine Datei mit 100 Farbnamen gemäß ihrer Verteilung im Bild (mind. 1%) geschrieben. Die Verteilungen der Farben über die beiden Bildkorpora hinweg zeigt einen exponentiellen Zusammenhang, so wie bei Wörtern in Textkorpora. Daher ist diese Vertextlichung von Bildeigenschaften geeignet für gängige Retrievalmodelle (siehe Abb. 3). CBIR eignet sich um menschlich verfasste textuelle Beschreibungen von Bildern zu ersetzen oder zu vereinfachen.[8] Das macht es interessant für den Kulturbereich, der über überproportional viele Bilddaten im Vergleich zu Mitarbeiter*innen verfügt. Unter den für CBIR im Kulturerbebereich nötigen Designentscheidungen ist für **Kusu** die Performanz

⁸ siehe dazu die [Europeana](#) und [iArt](#), sowie die [Suchmaschine des MKG](#), [digiporta](#) und [museum digital](#)

⁹ sofern Bilddaten vorlagen

relevant. [8] Um die Analysezeit zu verringern erfolgte die Farbboxtraktion auf zu einem Viertel der Originalgröße verkleinerten Bildern.

Interface-Design Das Front-End basiert auf HTML, CSS, und JavaScript. Neben einer klassischen Texteingabe für die Anfrage gibt es auswählbare Checkboxen für die Kategorien und für den farblichen Zugang zu den Bildern. Ein Button-Click löst die Weiterleitung der Anfragen über eine Form Request an eine intern laufende Flask-Instanz aus. Die User-Eingabe wird als Text in Python ausgelesen und an das Back-End übermittelt. Das Back-End gibt im Anschluss eine Liste an Resultaten zurück, die die Result-Tabelle ausfüllt (siehe Abb. 2).

4 Evaluation

User Gruppen Zur Bewertung der *Effektivität* **Kusus** nutzen wir Relevanz-Feedback für eine Sammlung an Queries. Die Queries stammen von Testpersonen, die, analog zu den vom MET festgestellten Motivationsgruppen, in **Fachpersonen**, **Kunstschaffende**, und **Fachfremde** eingeteilt werden (siehe Tabelle 2).

Variants Die Effektivitäten von acht **Variants** werden verglichen. Diese entstehen aus allen Kombinationen der Parameter **Ähnlichkeitsmaß**, **Analyzer**, sowie **Query-Mode**. Ähnlichkeitsmaße sind *BM25* und *Boolean*, Analyzer die beiden gängigen deutschen Elasticsearch Stemmer *german_stemmer* und *german_light_stemmer*, gewrappt durch die Analyzer *german_analyzer* und *german_light_analyzer*. Query-Modes sind die disjunktive Multiterm-Query (bezeichnet als *only_disjunction*) sowie eine Kombination aus Phrase-Match mit Gewicht 2, Konjunktion mit Gewicht 2, und Disjunktion mit Gewicht 1 (bezeichnet als *combined_op*).

Zur Messung der Effektivität der Variants werden die Maße *Precision@k* und *nDCG@k*, für $k \in \{5, 10, 20\}$ genutzt.

Relevanz Feedback Für jede Gruppe liegen 10 Anfragen vor. Für jede Anfrage poolen wir die Ergebnisse der jeweils top $k = 20$ Ergebnisse jedes Variants, präsentieren sie in zufälliger Reihenfolge um einen Ranking Bias zu vermeiden und bewerten sie auf ihre Relevanz, von 0 bis 5 (0 = nicht relevant, 5 = sehr relevant). Bei der Relevanzeinschätzung hilft das von den Testpersonen beschriebene Information Need zur jeweiligen Query.

Ergebnisse Abb. 4, 5 und 6 zeigen, dass für alle k -Werte und alle Variants der *nDCG* bei der User-Gruppe der Fachfremden (Laien) am niedrigsten ist. Unter den 8 Variants ist der Query-Mode der einzige Parameter, der sichtlichen Einfluss auf den *nDCG@k* hat (*combined_op/only_disjunction*).

Außerdem erzielt der Query-Mode *combined_op* signifikant bessere *Precision@10* Werte, als *only_disjunction*; es besteht kein signifikanter Unterschied zwischen den beiden Analyzern oder den beiden Ähnlichkeitsmaßen (siehe Tabelle 3).

In der Gruppe der Kuntschaffenden ist *combined_op* gegenüber *only_disjunction* erneut überlegen. Es erzielt deutlich bessere Precision@k Werte für $k = 5$ sowie $k = 20$ (siehe Abb. 7).

Es ist insgesamt zu erkennen, dass weder die beiden Stemmer, noch die Ähnlichkeitsmaße einen signifikanten Einfluss auf die Parameter Precision@k oder nDCG haben. Eine signifikante Verbesserung der eben genannten Parameter erfolgte jedoch durch die Kombination aus Phrase-Match, Konjunktion und Disjunktion mit verschiedenen Gewichten.

5 Fazit und Ausblick

Beim Vergleich des Einflusses verschiedener Index-Konfigurationen auf die Effektivität des Retrieval stellt sich heraus, dass diese keinen signifikanten Einfluss auf die Effektivität haben. Dies könnte an der mangelnden Verbosität der Daten liegen. Im Gegensatz dazu hat die Query-Verarbeitung einen signifikanten Einfluss. Hier bleibt offen wie ein eventueller Bias unseres eigenen Teams das Relevance Judgement beeinflusst haben könnte. Jedes Relevance Feedback wurde von genau einer Person durchgeführt. Eine Verbesserung in dieser Hinsicht wäre, das Relevance Feedback für jedes Ergebnis von mindestens zwei Assessoren durchzuführen, um deren Reliabilität besser einschätzen zu können.

Die Nutzung von Retrieval-Technologien mit Museums-Daten erfordert spezifisches Domänenwissen, das unter den Nutzer*innen ungleich verteilt ist.

Kusu implementiert Ansätze zur Befriedigung verschiedener Nutztendenbedürfnisse und schafft Zugang zu unklaren Informationsbedürfnissen über *CBIR*.

Um den Zugang zu den Daten zu erleichtern, sollte **Kusu** zukünftig mit einem verbesserten Front-End, Anfragevervollständigung und automatischer Anfragegenerierung ausgestattet werden. Außerdem könnte das Retrieval durch integrierte semantische Konzepte (z.B. Arbeit mit Synonymen) voraussichtlich stark verbessert werden. Zur automatisierten Klassifizierung von Objekten könnten Methoden wie *Topic Modelling* dienen.

Anhang

A Tabellen

Tabelle 1: Datensätze

Datengeber	MKG	Lampensammlung	Touché DigiPorta	gesamt
Jahr	2017	2019	2019	
Objektanzahl	12527	254	6070	18851
Format	LIDO-XML	LIDO-XML, Bilder	LIDO-XML, Bilder	
Beschreibung	über CdV	über CdV	über CdV	
Zugang	über Github	über Nextcloud	über DMD Plus	
längstes Dok.	46795 Zeichen	3237 Zeichen	108937 Zeichen	108937 Zeichen
kürzestes Dok.	2267 Zeichen	1732 Zeichen	4496 Zeichen	1723 Zeichen
durchschnittl.	4991 Zeichen	2255 Zeichen	13853 Zeichen	7808 Zeichen

Tabelle 2: Einteilung der Testpersonen, vgl. [17]

Name	Beschreibung	Motivation MET
Fachpersonen	Personen, die mindestens einmal in einem professionell-akademischen Kontext Erfahrungen mit Museen, Objekten, deren Daten oder Suchmaschinen auf diesen Daten sammelten	Intellectual
Kunstschaffende	z.B. Künstler*innen und Designer*innen	Inspirational
Fachfremde	Vergleichsgruppe um den Umgang derer mit Kusu einzuschätzen, die kein Domänenwissen haben	Visit Planners

Tabelle 3: T-Test für Precision@10, Fachfremde – a1=german-analyzer, a2=german-light-analyzer, – s1=boolean, s2=bm25, – q1=only-disjunction, q2=combined-operators

	a2-s1-q2	a2-s1-q1	a2-s2-q1	a1-s2-q1	a1-s1-q2	a2-s2-q2	a1-s2-q2	a1-s1-q1
a2-s1-q2	0.0	0.347	0.347	0.347				0.347
a2-s1-q1	0.347	0.0			0.347	0.347	0.347	
a2-s2-q1	0.347		0.0		0.347	0.347	0.347	
a1-s2-q1	0.347			0.0	0.347	0.347	0.347	
a1-s1-q2		0.347	0.347	0.347	0.0			0.347
a2-s2-q2		0.347	0.347	0.347		0.0		0.347
a1-s2-q2		0.347	0.347	0.347			0.0	0.347
a1-s1-q1	0.347				0.347	0.347	0.347	0.0

disj = c1 bool=s1 ga = a1

B Beispieldaten

Beispiel 1: Beispiel für einen ein Objekt beschreibenden LIDO-XML-Eintrag von kult westmuensterland.

```
<lido:descriptiveNoteValue> Petroleumlampe. Glaslampe der
    französischen Manufaktur Clichy. Mit Klarglas zusammen
    geschmolzene farbige Glassplitter ergeben ein dem
    Millefiori ähnliches Glas. Rot-weiß geflammt.
    Achteckige Grundfläche der Vase. Unten und oben
    ausgebuchtet, im mittleren Bereich schmaler.
    8'''Kosmosbrenner. Dochtrad mit Sonnenrelief.
    Brennerhöhe: 8 cm
</lido:descriptiveNoteValue>
```

C Bilder

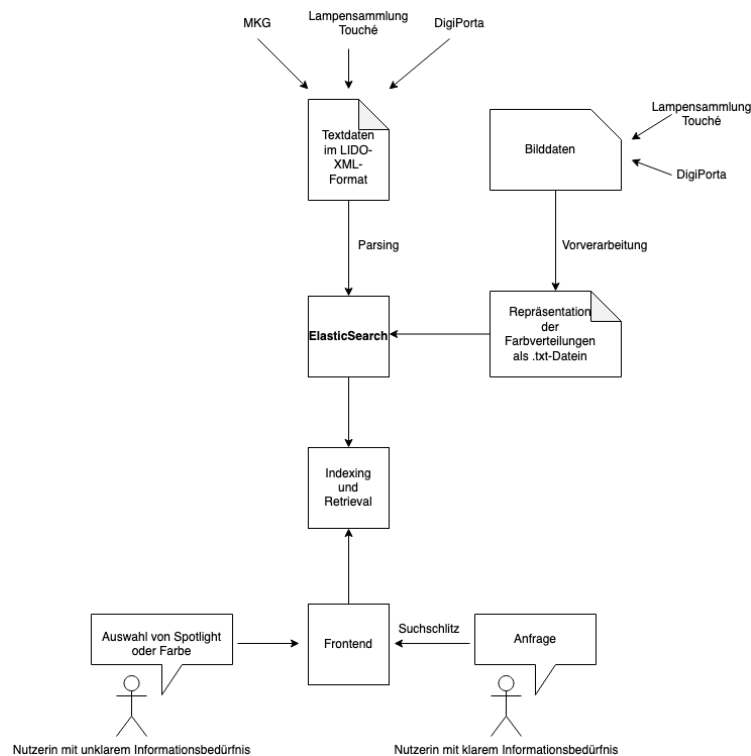


Abb. 1: Aufbau der Suchmaschine

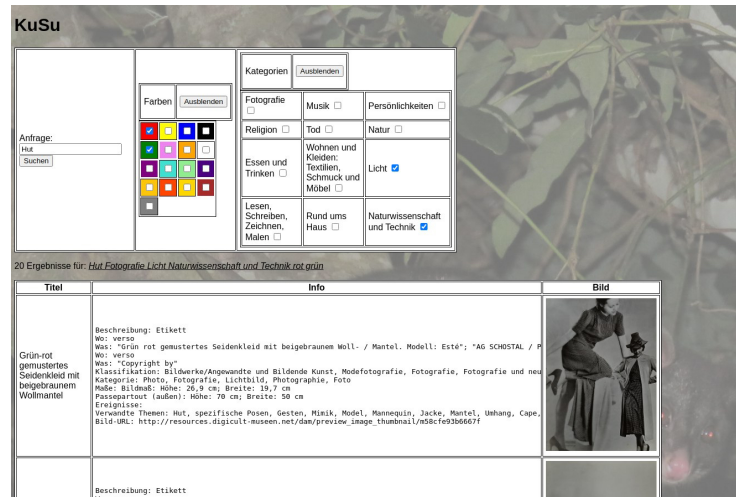


Abb. 2: Frontend der Suchmaschine

D Diagramme

Bildanalyse

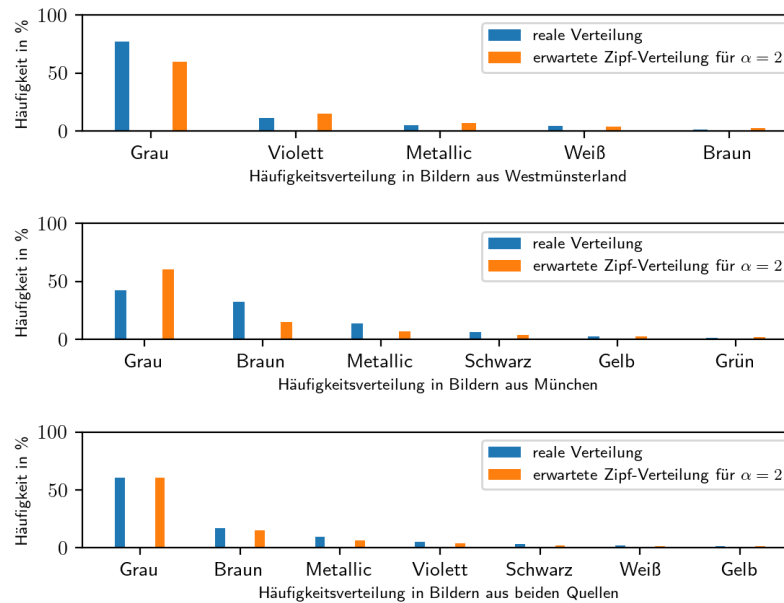


Abb. 3: Bildanalyse: Häufigkeitsverteilungen

Evaluation

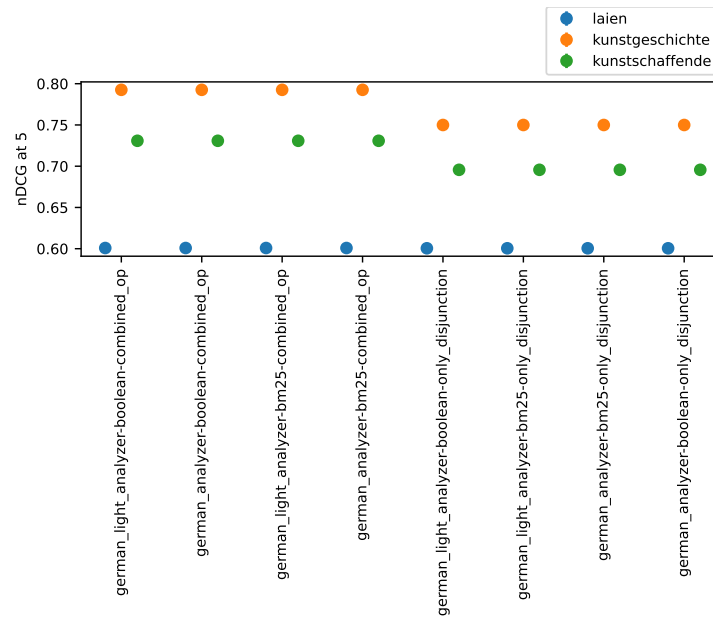


Abb. 4: nDCG at 5

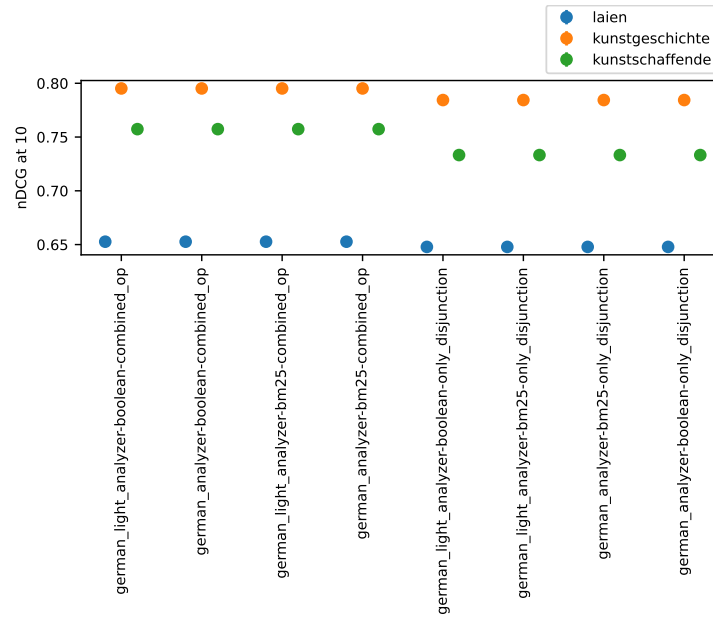


Abb. 5: nDCG at 10

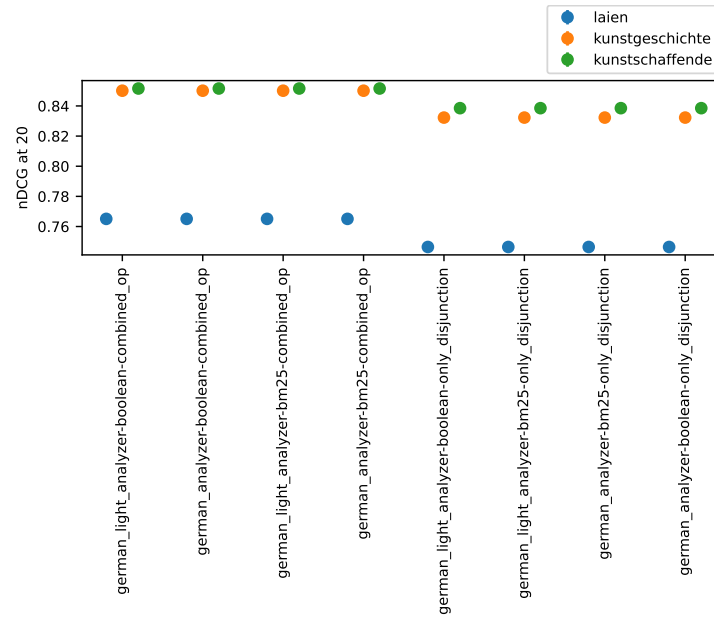


Abb. 6: nDCG at 20

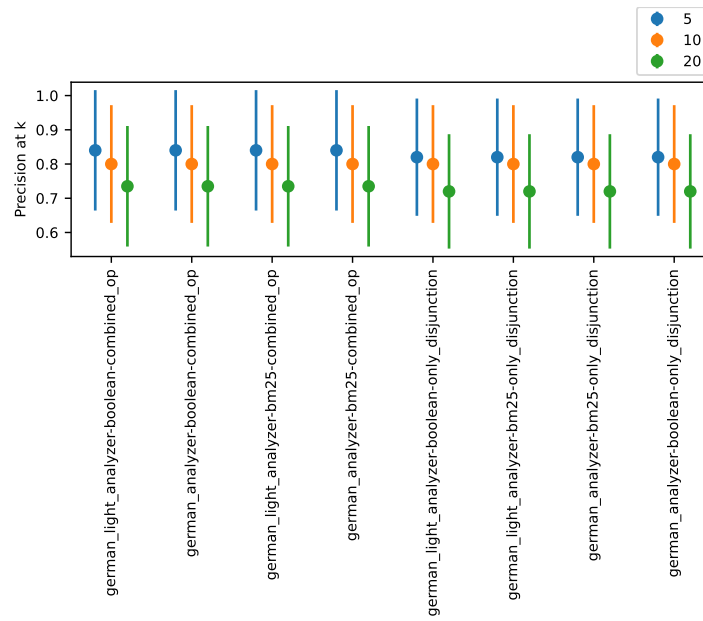


Abb. 7: Precision at k, Kunstschaftende

Literatur

1. Buttcher, S., Clarke, C.L., Cormack, G.V.: Information retrieval: Implementing and evaluating search engines. MIT Press (2016)
2. Dörk, M., Carpendale, S., Williamson, C.: The information flaneur: A fresh look at information seeking. In: Proceedings of the SIGCHI conference on human factors in computing systems. pp. 1215–1224 (2011)
3. Gormley, C., Tong, Z.: Elasticsearch: the definitive guide: a distributed real-time search and analytics engine. Ö'Reilly Media, Inc.”(2015)
4. Haffner, D.: Datenausgabebedarfe, Open Access. In: Handbuch Museum, pp. 194–198. Springer (2016)
5. Harman, D., et al.: Information retrieval: the early years. Foundations and Trends® in Information Retrieval **13**(5), 425–577 (2019)
6. Kohle, H.: Museen digital: Eine Gedächtnisinstitution sucht den Anschluss an die Zukunft. Heidelberg University Publishing (heiUP) (2018)
7. Petras, V., Hill, T., Stiller, J., Gäde, M.: Europeana—a search engine for digitised cultural heritage material. Datenbank-Spektrum **17**(1), 41–46 (2017)
8. Purificato, E., Rinaldi, A.M.: Multimedia and geographic data integration for cultural heritage information retrieval. Multimedia Tools and Applications **77**(20), 27447–27469 (2018)
9. Ranjgar, B., Sadeghi-Niraraki, A., Shakeri, M.: Cultural Heritage Information Retrieval: Data Modelling and Applications (2019)
10. Schmidt, A.: Warum »Access« nicht genug ist. Die MKG Sammlung Online und das Potenzial von offenen und nachnutzbaren Sammlungen. Konferenzband EVA, Berlin (2016)
11. Schmidt, A.: Voraussetzungen für und Herausforderungen von Co-Creation mit digitalen Museumssammlungen. In: Kultur in Interaktion, pp. 51–62. Springer (2020)
12. Schmidt, A.: persönliche Kommunikation (2022)
13. Springstein, M., Schneider, S., Rahnama, J., Hüllermeier, E., Kohle, H., Ewerth, R.: iART: A Search Engine for Art-Historical Images to Support Research in the Humanities. In: Proceedings of the 29th ACM International Conference on Multimedia. pp. 2801–2803 (2021)
14. Sugimoto, C.R., McCain, K.W.: Visualizing changes over time: A history of information retrieval through the lens of descriptor tri-occurrence mapping. Journal of Information Science **36**(4), 481–493 (2010)
15. Terras, M.: So you want to reuse digital heritage content in a creative context? Good luck with that. Art Libraries Journal **40**(4), 33–37 (2015)
16. Villaespesa, E.: Who Are the Users of The Met’s Online Collection? (2017)
17. Villaespesa, E.: Museum collections and online users: development of a segmentation model for the metropolitan museum of art. Visitor Studies **22**(2), 233–252 (2019)
18. Zindel, Y.: Zukunftsvision Museum: Institutionen und das Digitale. <https://19.re-publica.com/de/session/zukunftsvision-museum-institutionen-digitale> (05 2019)