

Gregor Neuböck

Zeitgemäße visuelle Darstellung von Digitalisaten in einem Repository – am Beispiel der DLOÖ – kann nur gelingen, wenn Daten geeignet in Format gebracht werden

Handbuch Repositorienmanagement, Hg. v. Blumesberger et al., 2024, S. 467–487
<https://doi.org/10.25364/978390337423225>



Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung 4.0 International Lizenz,
ausgenommen von dieser Lizenz sind Abbildungen, Screenshots und Logos.

Gregor Neuböck, Wienbibliothek im Rathaus, gregor.neuboeck@wienbibliothek.at | ORCID iD: 0000-0002-2979-4356

Zusammenfassung

Auf Basis der rasanten Entwicklungen im Bereich der Informationstechnologien konnten die Mittel zur Darstellung von Büchern in digitaler Form einen unglaublichen Qualitätssprung machen. Grundlage dafür ist die Qualität der Datensätze und Scans, wobei letztere sehr stark von der Qualität der Scanner, den Einstellungen und den Arbeitsabläufen abhängt. Die Qualität der Datensätze wiederum hängt von verschiedenen weiteren Faktoren ab. Die Orientierung an internationalen Standards kann dabei sicherlich zu den wichtigsten Kriterien gezählt werden, aber auch die Granularität der erfassten Daten stellt eine essentielle Grundlage für eine optimierte Visualisierung dar. Im Anschluss erläutere ich die derzeitigen Möglichkeiten der Visualisierung anhand der Digitalen Landesbibliothek Oberösterreich DLOÖ. Der hier vorliegende Beitrag behandelt genau diese Faktoren möglichst umfassend, sowie im Kernpunkt die darauf aufbauenden visuellen Darstellungsmöglichkeiten, ohne den Rahmen dieses Beitrags zu sprengen.

Schlagwörter: Visualisierung; Digitalisierung; Repotorium; Goobi; Retrodigitalisierung

Abstract

The Contemporary Visual Representation of Digital Copies in a Repository – Using the DLOÖ as an Example – Can Only Succeed if the Data Are Formatted Appropriately

The developments in the field of information technologies in the recent years have been so rapid that today the means of presenting books in digital form have improved in an incredible way, thus ensuring advancements in the quality of the digitisation. The prerequisite for this is the quality of the data sets and the scans, the latter depending very much on the quality of the scanners, their settings and the workflows in scanning mode. Besides, the quality of the records depends on several other factors. The orientation to international standards can certainly be counted among the most important criteria, but the granularity of the collected data is also an important basis for optimised visualizations. In the following, I will explain the current possibilities of visualisation by reference to the Digital State Library of Upper Austria DLOÖ. This contribution deals with precisely the above-mentioned factors as comprehensively as possible as well as with the resulting visual display options, without going beyond the scope of the publication.

Keywords: Visualisation; digitisation; repository; Goobi; retro-digitisation

1. Einleitung

Die Digital Humanities treiben die Entwicklungen zur Visualisierung von Daten seit vielen Jahren mit zunehmendem Tempo voran.

Visualisierung kann dabei helfen, Daten möglichst umfassend, aber gleichzeitig klar und übersichtlich darzustellen, ohne dabei wichtige Details zu verbergen oder gar zu verändern. Das Gesehene in einem Repository ist niemals mit einer wirklichkeitsgetreuen Abbildung gleichzusetzen, sondern ist immer eine Folge der Verarbeitung des einzelnen Individuums auf Basis dessen Vorwissens bzw. Wissensstands¹. In diesem Sinne ist bei der Gestaltung von Visualisierungen darauf zu achten, dass diese möglichst selbsterklärend sind und im besten Fall eine Art lernende Oberfläche² bilden. Im Idealfall werden die Benutzer:innen einerseits bei ihrem individuellen Wissensstand abgeholt und andererseits sorgen die Visualisierungen für eine optimale Wissensorganisation. Die Digitalisierung bzw. die Entwicklung des Semantic Web hat dazu geführt, dass nicht nur die Objekte selber (Scans der einzelnen Seiten) sondern auch ihre Beschreibungen digital verfügbar sind. Es gibt also einen rasant wachsenden Bereich an Struktur- und Metadaten. Damit in diesem riesigen Datenmeer der Überblick gewahrt bleiben kann, Suchanfragen also erfolgreich und zielgerichtet, aber niemals unwissentlich ausschließend sind, ist es zielführend, die Wissensorganisation über Modellansätze zu realisieren. Dazu zählen z. B. die Taxonomie, die in abgewandelter Form heute auch in Form des Tagging angewendet werden kann, Thesauri im Sinne einer Vernetzung verwandter Begriffe (siehe z. B. Named Entity Recognition, NER) in einem semantischen Netz sowie Ontologien, die man als³ „[...] logisch einwandfrei definierte, maschinell interpretierbare Beschreibungen des Weltwissens [...]“⁴ bezeichnen kann. NER bezieht sich auf die automatisierte Erkennung von Eigennamen durch Algorithmen, was wesentlich schwieriger ist, als es auf den ersten Blick erscheint. Denn Entitäten sind nicht immer sofort ganz klar und eindeutig zu erkennen. Orts-, Personen-, Organisations- und Produktnamen werden ganz eindeutig als Entitäten bezeichnet.⁵

1 Pamperl, B. (2017), S. 90f.

2 Im Sinne einer *Lernenden Organisation* sollte sich auch eine Lernende Oberfläche den Benutzer:innen anpassen, indem z. B. Inhalte, in Abhängigkeit der Nutzung, einem Ranking unterzogen werden sollten.

3 Keller, S. et. al. (2014), S. 11–14.

4 Keller, S. et. al. (2014), S. 13.

5 Roth, J. (2002), S. 5f.

Paik et al.⁶ unterscheiden neun Hauptkategorien. Eine Kategorie wird als Misc.⁷ bezeichnet. Alle Hauptkategorien besitzen zudem Unterkategorien, wodurch sich schlussendlich 30 Kategorien ergeben.

In der Retrodigitalisierung steht verstärkt das Bild im Zentrum und wird, so wie Pamperl es beschreibt, zum Wissenschaftsbild⁸. Windhager spricht in der Biographieforschung von einer „Weiterentwicklung von Bildern zu genuinen Funktionsträgern eines dynamischen und kritischen geschichtswissenschaftlichen Diskurses“⁹.

Die reale Umsetzung von Digitalisierungsprojekten erweist sich immer als hochkomplexer, extrem herausfordernder und permanenter Erneuerungsprozess.

Allen Beteiligten sollte von Anfang an klar sein, dass Digitalisierungsprojekte einen langen Atem sowie personelle und monetäre Ressourcen erfordern, die budgetiert und bereitgestellt werden müssen. Im Vorfeld sollte also schon abgeklärt werden, ob diese von den jeweiligen Institutionen dauerhaft aufgebracht werden können.

Neben den Bibliothekar:innen und Programmierer:innen kann man nur dazu raten, pädagogisch geschultes Personal und Anwender:innen in Entwicklungsprozesse einzubinden, damit nicht an den Benutzer:innen vorbeientwickelt wird.

Der hier vorliegende Beitrag beschäftigt sich mit zentralen Komponenten von Repositorien, welche die Visualisierung von Daten und die damit verbundene Qualität der Darstellung für die Benutzer:innen erheblich beeinflussen. Da auch an der Oberösterreichischen Landesbibliothek die Anzahl der digitalisierten Werke¹⁰ und in einem noch wesentlich höheren Ausmaß, die mit diesen Werken verbundenen Struktur- und Metadaten eine nicht mehr überblickbare Menge erreicht haben, sind auch wir auf verschiedene Werkzeuge der Digital Humanities zur Informationsvisualisierung angewiesen. Etwa, wie man z. B. Suchanfragen oder vorhandene Geodaten geeignet darstellt.¹¹ Auf Basis dieser Daten werden die derzeit verwendeten Werkzeuge und deren Visualisierung innerhalb der Digitalen Landesbibliothek Oberösterreich (DLOÖ) vorgestellt.

⁶ Paik, L. et al. (1993), S. 156.

⁷ Unter dieser Kategorie werden alle Entitäten angeführt, die keiner der anderen Kategorien zugeordnet werden können.

⁸ Pamperl, B. (2017), S. 92.

⁹ Windhager, F. (2017), S. 71.

¹⁰ Mit Stand vom 01.10.2021 sind innerhalb der Digitalen Landesbibliothek Oberösterreich (DLOÖ <https://digi.landesbibliothek.at>) beinahe 6.500 Werke digitalisiert, was ca. 600.000 Seiten entspricht.

¹¹ Pamperl, B. (2017), S. 93f.

Es besteht weder der Anspruch auf Vollständigkeit aller einflussnehmenden Faktoren, noch auf detaillierte Ausarbeitung dieser. Vielmehr geht es darum, einen praxisnahen Blick auf diese zu werfen und Anregungen für die eigene Arbeit zum Thema Visualisierung in der Digitalisierung zu finden.

2. Qualität der Datensätze

2.1. Standards

Wie eingangs erwähnt, hängt die Qualität der Datensätze sehr stark von der Orientierung an internationalen Standards ab¹². Bei der Datenvisualisierung¹³ sollte man sich immer bewusst machen, dass diese schon eine Form der Interpretation von Rohdaten darstellt. In diesem Sinne sollte das Ergebnis immer kritisch hinterfragt werden.¹⁴

Auch kann es im Laufe der Zeit passieren, dass sich die Datenfelder vielmehr an den Wünschen der einzelnen Protagonist:innen¹⁵ als an internationalen Standards orientieren. Verlockend erscheinen die grenzenlosen Möglichkeiten, den Aufbau von Datenfeldern abseits von Normierungen ausschließlich an den Wünschen der Beteiligten zu orientieren.

Man sollte sich dessen bewusst sein, dass das Erstellen von Datenfeldern, ohne sich über deren Normierung Gedanken zu machen, zu datenbanktechnischen Problemen mit langfristigen Auswirkungen führen kann. Repositorien erfordern zwingend laufend Verbesserungen und müssen an aktuelle technische Standards angepasst werden¹⁶. Werden bei diesen Erneuerungs- und Entwicklungsprozessen Standards nicht eingehalten, ist mittel- und langfristig ein erheblicher Mehraufwand für die Erhaltung und Pflege der Daten zu erwarten. Im schlechtesten Fall kann dies sogar zu Datenverlust führen, weil keine maschinelle Verarbeitung mehr möglich ist.

Zudem gewinnt die Normierung von Datenfeldern noch rascher an Bedeutung, wenn das eigene Repository an internationale Plattformen (Europeana¹⁷, Verzeichnis digitalisierter Drucke¹⁸,....) angedockt werden soll. Dann wird es nötig, eine

12 <http://www.loc.gov/standards/mets/>

13 z. B. Suchanfragen, Timelines, Kalendersuche etc.

14 Pamperl, B. (2017), S. 95.

15 Wissenschaftler:innen, Bibliothekar:innen, Sammlungsleiter:innen und dgl.

16 <https://forschungsdaten.info/themen/beschreiben-und-dokumentieren/metadaten-und-metadatenstandards/>

17 <https://www.europeana.eu/de>

18 <https://www.zvdd.de/startseite/>

Nachnutzung über APIs, Schnittstellen (z. B. IIIF¹⁹, REST, OAI-PMH²⁰, SRU) und anerkannte Datenstandards wie z. B. JSON zu ermöglichen. In der DLOÖ werden viele verschiedene maschinell lesbare und international anerkannte Datenformate via OAI-PMH angeboten.

Der Wissenschaftsbetrieb ist bestrebt, internationale Datenstandards einzuhalten, damit ein möglichst reibungsloser und barrierefreier Datenaustausch ermöglicht wird. Dieser kann zusätzlich noch verbessert werden, indem auf Titeläquivalenz²¹ geachtet wird, Farbe und Kontrast parallel eingesetzt werden, Optionen zur Tastaturnavigation angeboten werden, strukturierte Inhalte vorhanden sind, die Verständlichkeit²² gegeben ist und eine Geräteunabhängigkeit²³ garantiert werden kann.²⁴

2.2. Scans

Die Qualität der Scans ist ein maßgebender Faktor bei der Visualisierung von Daten. Auch im Sinne eines bestandsschonenden Umgangs sollte auf eine nachhaltige Scanstrategie geachtet werden, um nicht alle paar Jahre in die Verlegenheit zu geraten, Scans neu anlegen zu müssen. Es liegt auf der Hand, dass jede noch so schonende Prozedur Bücher belastet, was angesichts des Mehrwerts, Materialien unabhängig von Ort und Zeit für Wissenschaft und Allgemeinheit verfügbar zu machen, in Kauf genommen wird.

Beim Neukauf eines Scanners sollte man sich zu Beginn einen Überblick über geeignete Scangeräte hinsichtlich Scangröße, Scanauflösung und Farbtiefe verschaffen. Die Scanauflösung Dots per Inch (DPI) ist eine drucktechnische Angabe und nutzt das Farbmodell CMYK. Pixel per Inch (PPI) ist eine digitale Angabe und verwendet das additive Farbmodell RGB. Die Farbtiefe ist ein entscheidender Wert, um eine wirklichkeitsgetreue Abbildung von Farben zu ermöglichen. Lange Zeit galt eine Farbtiefe von 24 Bit (True Color) als der beste Standard, mittlerweile werden auch Systeme mit bis zu 30 Bit (Deep Color) bei Scannern angeboten.

Neben diesen grundlegenden Parametern sollte man sich auch überlegen, ob man einen Zeilenscanner oder einen Scanner mit einem Flächenchip (vergleichbar mit

19 <https://docs.goobi.io/goobi-viewer-de/conf/1/33/2>

20 <https://digi.landesbibliothek.at/viewer/oai>

21 Jede Grafik benötigt eine Textbeschreibung.

22 Richtige Sprachauszeichnung, Verständlichkeit der Navigation, angebotene Orientierungshilfen.

23 Verschiedene Betriebssysteme, Unterstützung verschiedener Versionierungen von Softwareprodukten.

24 Kogler, G. (2014), S. 5–8.

einer Digitalkamera) einsetzen möchte. Zeilenscanner bieten die besseren Auflösungen über große Flächen und arbeiten mit einem Zeilensor (Abtastrate beachten!), der Zeile für Zeile das darunterliegende Objekt ausliest und gleichzeitig für jedes Pixel alle RGB-Farben aufnimmt. Kamerasysteme sind ungleich flexibler und weisen eine wesentlich höhere Scangeschwindigkeit auf. Allerdings müssen durch ihren spezifischen Bildsensor, Bildpunkte im Blau- und Rotbereich stark interpoliert (digital ergänzt) werden.

Digitalisierung ist immer im Dilemma zwischen den Ansprüchen, möglichst nahe am Original zu sein und gleichzeitig einen optimierten Digitalisierungsworflow zu etablieren. Um die Auflösung eines Scanners beurteilen zu können, ist es wichtig, zwischen absoluter²⁵ und relativer²⁶ Auflösung zu unterscheiden. So wird vielleicht auch klar, dass eine hohe relative Auflösung per se noch kein scharfes Bild garantiert. Wenn z. B. in das Bild stark reingezoomt wird, sinkt im Bildausschnitt die relative Auflösung, bis das Bild irgendwann unscharf wirkt. Beim Scanner-Ankauf ist daher darauf zu achten, über welche Fläche eine optische Auflösung garantiert wird, um beim Zoomen auch noch scharfe Bilder zu ermöglichen. An der Oberösterreichischen Landesbibliothek wird jedes Objekt seit über zehn Jahren mit einer relativen Auflösung von 600 PPI gescannt. Nur ganz selten verringern wir bei sehr großen Karten oder Leporellos die Auflösung, um nicht allzu große Bilddateien zu erhalten.

Hat nun eine engere Auswahl an geeigneten Geräten stattgefunden, sollten alle in Betracht kommenden Scanner mit eigenen Materialien²⁷ eingehend getestet werden. Neben Problemen bei den mechanischen Abläufen²⁸ kann es auch vorkommen, dass Scanner die Farben nicht wirklichkeitsgetreu abbilden oder die angegebene Auflösung²⁹ gar nicht erreichen bzw. mit interpolierten Auflösungen³⁰ arbeiten. Auch die Scangeschwindigkeit sollte in einem eigenen Test über verschiedene Flächen mit unterschiedlichen Auflösungen überprüft und mit den Herstellerangaben verglichen werden. Die Ausführung und Qualität der Linsen beeinflussen alle

²⁵ Entspricht der maximalen Scanfläche von z. B. 1250x3200 Pixel.

²⁶ Entspricht der Anzahl an Pixel/Dots per Inch, also einer Dichteangabe.

²⁷ Man sollte eine heterogene Auswahl aller in der Bibliothek vorkommenden Dokumententypen testen.

²⁸ Wie öffnet die Glasplatte? Gibt es eine Buchstütze? Wie wird das Buch eingelegt? Kann ich auch ohne Glasplatte scannen?...

²⁹ Sollte immer mit dem Zusatz „optisch“ angegeben sein, z. B. 600PPI optisch.

³⁰ Diese werden durch Algorithmen errechnet und sollten unbedingt vermieden werden.

Qualitätskriterien eines Bildes und sollten mit eigenen Materialien eingehend geprüft werden³¹. Testcharts für Auflösung³² und Kalibrierung³³ helfen dabei, die technischen Angaben normiert zu überprüfen und die Daten der Hersteller zu verifizieren. Regelmäßig sollte die Kalibrierung eines Scanners im Alltagsbetrieb vorgenommen werden, um abhängig von den jeweiligen Herstellerangaben optimale Ergebnisse und eine gleichbleibende Qualität zu halten.

3. Software

Welche Software man als Repository einsetzt, hängt stark von den benötigten Funktionalitäten und Anforderungen ab.

An der Oberösterreichischen Landesbibliothek kommt eine international verwendete Open-Source Software zum Einsatz, die von vielen verschiedenen Protagonisten entwickelt wird.

Noch vor wenigen Jahren stand man Open Source sehr kritisch gegenüber, da man der Ansicht war, dass die Weiterentwicklung dadurch nicht gewährleistet sei, ganz im Gegensatz zu proprietärer Software, der man genau diese Eigenschaft zusprach. Mittlerweile hat sich dieses Bild stark gewandelt, denn es hängt nicht von der Lizenz einer Software, sondern von der Bedeutung der Community ab, welche eine Software einsetzt. Gibt es wesentliche Stakeholder, kann davon ausgegangen werden, dass die Software langfristig weiterentwickelt wird und die Community über das entsprechende Entwicklungspotential verfügt. Einerseits ist dies wichtig, um Features zu entwickeln, die die aktuell vorhandenen technischen Möglichkeiten ausreizen, andererseits um einen möglichst barrierefreien Zugang auf allen möglichen Endgeräten mit verschiedensten Betriebssystemen auf unterschiedlichen Browsern³⁴ zu garantieren.

Ein weiteres Kennzeichen von Open Source ist der permanente Entwicklungsprozess, was nichts mit dem in der Softwareentwicklung verwendeten Begriff der Betaversion³⁵ zu tun hat, sondern auf einen andauernden Entwicklungsprozess hinweist, der wiederum der Garant für Innovation, Stabilität und Aktualität ist. Die Schlussfolgerung daraus muss sein, dass man als Projektmanager:in Softwareentwicklung als einen permanenten Verbesserungsprozess sieht. Nichts ist weniger

31 Gold kann einen grünlichen Anteil bekommen oder es kommt im Randbereich von Bildern wegen größerer Linsenfehler zu Verzerrungen.

32 USAF Resolution Test Chart

33 ISO-Standards 12641-2

34 Auch alle aktuellen Versionierungen eines Browsers sollten fehlerfrei funktionieren.

35 Software, die noch nicht für den Produktionsbereich geeignet ist, weil noch zu viele Bugs vorhanden sind.

nachhaltig, als isolierte Digitalisierungsprojekte, die ein Budget und einen zeitlich begrenzten Ablauf haben, ohne dass in irgendeiner Form darüber nachgedacht wurde, wie eine dauerhafte Weiterführung oder Einbindung in vorhandene Digitalisierungsprojekte gewährleistet werden könnte. Die Gefahr, dass finanzielle und personelle Ressourcen nicht nachhaltig eingesetzt werden, ist hierbei groß.

In jedem Fall darf nochmals explizit darauf hingewiesen werden, sich an internationale Datenstandards zu halten, insbesondere auch darauf, dass die verwendete Software die Einhaltung dieser ermöglicht, denn nur sie garantieren am ehesten einen problemlosen Umstieg auf ein anderes Softwareprodukt, sollte dies einmal erforderlich sein.

Ein weiterer Punkt ist die Plattformunabhängigkeit der Software, sodass kein eigener Client für den Zugriff auf die Produktionssoftware erforderlich ist. Für den Zugriff sollten nur ein herkömmlicher Browser und eine Internetverbindung erforderlich sein. Serverbasierte Software kann also unabhängig von Ort, Zeit und Gerätekonfiguration verwendet werden. Nebenbei fallen bei Open Source auch keine Lizenzgebühren an und es können nach Bedarf neue Benutzer:innen angelegt werden. So ergibt sich kein großer Aufwand, externe Dienstleister, aber auch Mitarbeiter:innen im Homeoffice in den Arbeitsprozess einzubinden. Heute sind derartige Systeme in aller Munde und werden gemeinhin unter dem Sammelbegriff Cloud-Computing geführt.

4. Visualisierungen

Im Kapitel Visualisierungen werden verschiedene Anwendungsbereiche innerhalb der DLOÖ behandelt. Neben der Bildanzeige zählen dazu aber auch die Darstellungen der Struktur- und Metadaten, bzw. welche Anwendungsmöglichkeiten sich aus diesen Daten für die visuelle Darstellung in einem Repository allgemein ergeben.

Auch wenn im Zentrum dieses Beitrags die bildliche Darstellung der Digitalisate steht, bedarf dieses Themas doch einer viel umfassenderen Behandlung, weswegen auch auf die Performanz eingegangen wird, da diese als sehr entscheidend für hohe Kund:innenzufriedenheit angesehen werden muss. Die Konfiguration hat auf die Performanz einen großen Einfluss, daher wird auch diese in einem Unterkapitel angesprochen. Die meisten Projektmanager:innen an Bibliotheken werden wenig bis gar nichts an der Konfiguration verändern, nichtsdestotrotz ist es doch unumgänglich, dass man sich dieser von mir angesprochenen Parameter bewusst wird, damit man darauf Einfluss nehmen kann.

4.1. Einstellungen der Bildanzeige

4.1.1. Performanz

Da sich die ursprünglich gescannten Bildformate nicht für performante Darstellungen eignen³⁶, müssen diese komprimiert werden. An der Oberösterreichischen Landesbibliothek verwenden wir als Bildformat TIFF.³⁷ Diese Bilder dienen in erster Linie der Archivierung.

Zum Einsatz kommen verschiedene Kompressionsverfahren, wobei die JPEG-Komprimierung sicherlich eine der häufigsten Formen ist. Die eingesetzte Auflösung und das Format der Bilder hängen stark mit den jeweiligen Anforderungen im Repository zusammen. Eine von der Gerätekonfiguration bestimmte (Barrierefreiheit, Software-Ergonomie beachten!) Auslieferung der Bildformate in Abhängigkeit von Bildschirmgröße, Auflösung und Browser sollte im Idealfall automatisch ablaufen (Responsives Webdesign).

Die Einstellungen der Zoomfunktion können die Performanz erheblich beeinflussen. Möglichst hohe Zoomstufen sind aus Benutzer:innensicht sicherlich positiv zu bewerten, es sollte allerdings bedacht werden, dass die Bildgröße den Arbeitsspeicher erheblich belasten kann und als Folge davon die Performanz darunter leidet. Geeignete technische Lösungen (z. B. Kompressionsformate, Kacheln...) helfen, einen guten Kompromiss zwischen Qualität und Performanz zu finden, und tragen dazu bei, dass auch Zugriffe von leistungsschwächeren Geräten performant ablaufen können.

Eine weitere Möglichkeit wäre das Anbieten verschiedener Bildformate, damit man in Abhängigkeit von Internetverbindung und Arbeitsspeicher ein stets performantes System etabliert.

Eine begrenzte Anzahl der mit einem bestimmten Lademechanismus geladenen Bilder kann Speicherüberläufe bei sehr großformatigen Werken verhindern. Clientseitiges Proxy-Caching kann im Fall unerwünschter Effekte auf die Performanz unterbunden werden. Und schlussendlich darf ich noch auf den Unterschied zwischen server- und clientseitiger Rechenleistung hinweisen. Über die Softwarekonfiguration kann darauf Einfluss genommen werden.

³⁶ Bei den Scans handelt es sich um hochauflösende Bilder, die daher mit hohen Dateigrößen aufwarten.

³⁷ TIFF bietet eine verlustfreie Speicherung der Daten.

4.1.2. Konfiguration der Bildanzeige

Gibt es Probleme in der Bildanzeige, die auf lange Ladezeiten oder auf ein bestimmtes Softwareprodukt zurückzuführen sind, können diese bei Benutzer:innen zu großem Unmut führen. Im schlimmsten Fall führt es sogar dazu, dass diese als Kund:innen verloren gehen. Ein niedrigschwelliger Zugang zu einem Feedbacksystem ist eine Möglichkeit, diesem Unmut zu begegnen.

Ebenso sollte an allen nur möglichen Schrauben der Konfiguration gedreht werden, um ein schlankes und performantes System zu etablieren.

Das Kacheln³⁸ und die damit verbundene Bildgröße der in den verschiedenen Zoomstufen verwendeten Bilder stellt eine zentrale Möglichkeit zur Verbesserung der Performanz dar. Mit der Skalierung, den sogenannten ScaleFactors³⁹, werden die Kachelgrößen definiert, um so die verschiedenen Zoomstufen anzeigen zu können. Dies ist grundsätzlich sehr positiv zu sehen, weil dadurch Bilder wunschgemäß skaliert werden können, allerdings kann das bei sehr groß skalierten Bildern zu einer Überlastung des Arbeitsspeichers führen.

In Goobi wird zur Bildanzeige der webbasierte Viewer OpenSeadragon⁴⁰, basierend auf einer IIIF-API⁴¹, verwendet. Dieser lässt sich auf vielfache Weise bezüglich der zuvor genannten Parameter konfigurieren. Innerhalb der IIIF-API unterscheidet man Image API, Presentation API, Authentication API, Content Search API, Change Discovery API und Content State API.⁴² Diese Schnittstellen ermöglichen die Darstellung digitalisierter Inhalte auf unterschiedlichsten Systemumgebungen. Hinter dieser Initiative stehen viele Bibliotheken und Universitäten, aber auch Anbieter digitaler Inhalte.⁴³

4.2. Visualisierungen in Goobi

In Goobi⁴⁴ gibt es eine Fülle an verschiedenen Visualisierungen. Innerhalb der DLOÖ werden nicht alle Möglichkeiten, die Goobi zur Visualisierung anbietet, aus-

³⁸ Darunter versteht man das mosaikartige Zusammenbauen eines Bildes. Gleichzeitig wird durch diesen Mechanismus beim Aufbau eines großen Bildes wesentlich weniger Arbeitsspeicher benötigt.

³⁹ Goobi viewer Handbuch. Konfiguration der Bildanzeige <https://docs.goobi.io/goobi-viewer-de/conf/1/11/3>

⁴⁰ OpenSeadragon. 2021 <https://openseadragon.github.io/>

⁴¹ <https://docs.goobi.io/goobi-viewer-de/conf/1/33/2>

⁴² <https://iiif.io/>

⁴³ <https://iiif.io/community/consortium/members/>

⁴⁴ Softwarepaket für Digitalisierungsprojekte, bestehend aus Präsentations- und Produktionssoftware.

geschöpft. Auf den folgenden Seiten stelle ich die bei uns verwendeten Visualisierungen vor. Grundsätzlich sollte darauf geachtet werden, dass alle Elemente in einem Repository barrierefrei sind und auf unterschiedlichsten Geräten und Konfigurationen laufen. Um dies bestmöglich zu erreichen, ist eine Orientierung an den Richtlinien des W3-Consortiums⁴⁵ unumgänglich.

4.2.1. Bildanzeige

Die Bildnavigation sollte verschiedene Optionen enthalten, um den Benutzer:innen eine möglichst vielfältige, zugleich intuitive und einfache Bedienung zu ermöglichen. Neben einer Dropdown-Schaltfläche, um auf individuelle Buchseiten wechseln zu können, sowie den Blätterfunktionen „Vor“, „Zurück“, „erstes Bild“ und „letztes Bild“, ist eine Doppelseitenansicht und das Rotieren des Bildes in 90°-Schritten als obligatorisch anzusehen. Besonders möchte ich auf die Korrekturmöglichkeit (zugehörige Schaltfläche erscheint erst in der Doppelseitenansicht) der Doppelseitenansicht hinweisen, die immer dann erforderlich ist, wenn man sich beim Wechsel in die Doppelseitenansicht gerade auf einer Recto-Seite befunden hat, weil dann rechte und linke Seite des Buches genau vertauscht sind. Das Zoom kann entweder mit einem Schieberegler oder per Scrollrad der Maus stufenlos verändert werden.

Die Seitenvorschau dient dazu, Buchseiten mit Abbildungen oder anderweitigen Besonderheiten rasch identifizieren zu können. Da sie aber eine völlig andere Navigation als die der Einzelbilder benötigt, wird diese als eigener Menüpunkt links der Bildanzeige angezeigt⁴⁶. Größe und Auflösung der Vorschaubilder sind dabei so gewählt, dass vorhandene Abbildungen und dergleichen rasch erkannt werden können, ohne allerdings lange Ladezeiten zu verursachen.

Die Vollbildanzeige⁴⁷ ergänzt die normale Bildanzeige. Neben der Darstellung höherer Zoomstufen ist diese auch bei kleinen Bildschirmgrößen (Handy, Tablet) ein wichtiges Hilfsinstrument. In der Vollbildanzeige können mithilfe des Zooms Details von Büchern im Format Großfolio dargestellt werden, die beim Original ein gutes Vergrößerungsglas erfordern würden. Neben der Bildanzeige lassen sich bibliografische Daten, Werkzeuge zur Bildmanipulation oder der Volltext einer eventuell vorhandenen Transkription einblenden.

⁴⁵ <https://www.w3.org/>

⁴⁶ <https://digi.landesbibliothek.at/viewer/?thumbs/AC00969620/1/>

⁴⁷ <https://digi.landesbibliothek.at/viewer/fullscreen/AC05371191/265/>

Ein Dropdown-Feld zur Auswahl von Volltextseiten⁴⁸ sollte ebenfalls zur Verfügung gestellt werden. So können bei nur teilweise transkribierten Werken rasch die Seiten mit Volltext⁴⁹ herausgefiltert werden. Volltexte werden bei uns automatisiert für Antiqua- und Frakturschriften erstellt. Nicht jede Frakturschrift wird gut erkannt, insbesondere bei stark ausgeschmückten Frakturschriften und bei Versalien treten verstärkt Probleme auf, und bei den in Inkunabeln bevorzugt verwendeten Schriften wie z. B. Textura, Rotunda oder Breidenbach treten noch wesentlich größere Probleme in der Texterkennung auf. Das derzeit bei uns eingesetzte Produkt zur Texterkennung ist Tesseract, eine von Google entwickelte freie Software.

Nichtsdestotrotz bietet der Volltext Möglichkeiten, mit digitalisierten Büchern zu arbeiten, wie dies beim gedruckten Buch bisher nicht möglich war, und alternativ können fehlerhafte Volltexte mit einem Crowdsourcingmodul (siehe Kapitel 4.2.6.) verbessert werden.

Die Software Transkribus⁵⁰, welche schon von der UB Greifswald⁵¹ eingesetzt wird, stellt für alte Drucke und Handschriften eine wertvolle Ergänzung in der Texterkennung dar. Es handelt sich dabei um ein lernendes System zur automatisationsunterstützten Texterkennung. Das bedeutet, dass man die Software zu Beginn für eine bestimmte Schrift (gleich ob Druck- oder Handschrift) manuell trainieren muss. Mit jeder weiteren Trainingsseite werden stetig mehr Wörter eines Textes automatisch erkannt. Die Erkennungsraten können in der Folge auf deutlich über 95% steigen. Der Text kann anschließend in den Formaten PDF, Excel, DOCX, TXT und TEI exportiert und so in Goobi als Volltext eingespielt werden. Zurzeit gibt es schon unzählige gut trainierte Vorlagen für Druck- und Handschriften. Aktuell gibt es auch an der Oberösterreichischen Landesbibliothek Planungen, die Software als alternative OCR-Lösung in den Goobi-Workflow einzubauen.

4.2.2. Inhaltsverzeichnis

Den Menüpunkt Inhaltsverzeichnis⁵² möchte ich als Spezifikum unserer Bibliothek vorstellen. An der Oberösterreichischen Landesbibliothek wird jedes digitalisierte Buch mit einer möglichst hohen Dichte an Meta- und Strukturdaten versehen. Das bedeutet, wir erfassen alle Abbildungen, Kapitel, Vorwörter, Einleitungen, Regi-

⁴⁸ Dazu muss im Menü links der Bildanzeige zuerst von „Bildanzeige“ zu „Volltext“ gewechselt werden.

⁴⁹ <https://digi.landesbibliothek.at/viewer/fulltext/177/159/>

⁵⁰ Transkribus. READ-COOP SCE, 2021 <https://readcoop.eu/de/transkribus/>

⁵¹ https://www.digitale-bibliothek-mv.de/viewer/fulltext/PPNUAG-HGW_obj_5442087/1/

⁵² https://digi.landesbibliothek.at/viewer!/toc/AC02003800/1/LOG_0003/

ster, Tabellen usw. in Handarbeit. Große Player wie Google setzen zur automatisierten Verarbeitung von Meta- und Strukturdaten neuronale Netzwerke ein. Für kleinere Institutionen wie Bibliotheken und Archive sind derartige Anwendungen derzeit weder verfügbar noch leistbar.

In jedem Fall ist diese hohe Dichte an Meta- und Strukturdaten die Basis für vielseitige Visualisierungen bei Suchanfragen. Insgesamt spielen Visualisierungen bei komplexen Suchanfragen, spezifischen Sammlungen (Digitale Kollektionen⁵³), der Kalendersuche⁵⁴, der Zeitleiste⁵⁵ und der Darstellung von Geodaten innerhalb der DLOÖ eine tragende Rolle.

Die hohe Dichte an Meta- und Strukturdaten hat aber auch noch eine andere, sehr wesentliche Bedeutung für Bibliotheken. Google bewertet bei Suchanfragen die in Struktur- und Metadaten erfassten Einträge wesentlich höher als reinen Volltext und setzt diese dementsprechend im Ranking weiter nach vorne. Als Folge davon ergibt sich eine wesentlich verbesserte Außenwirkung der eigenen Institution, insbesondere was identitätsstiftende Bestände (Obderennsia) angeht.

4.2.3. Bibliografische Daten

Die bibliografischen Grunddaten eines Werkes⁵⁶ reichern wir mit vielen weiteren Daten an. Neben persistenten Verknüpfungen mit wichtigen Datenbanken, wie z. B. Hill Museum und Manuscripta speziell für Handschriften⁵⁷ und ISTC⁵⁸, GW⁵⁹ oder Hain⁶⁰ für Inkunabeln⁶¹, finden sich dort auch Lizenzangaben oder GND-Einträge, die zudem vollständig indexiert werden. So kann z. B. nach unterschiedlichen Schreibweisen eines Autors oder einer Autorin gesucht werden. Mit Göthe, Johann W. êvoné findet man in der Erweiterten Suche (nach Autor und Sammlung Handschriften) exakt einen Brief von ihm, der sich in unserem Besitz befindet.

⁵³ <https://digi.landesbibliothek.at/viewer/browse/>

⁵⁴ <https://digi.landesbibliothek.at/viewer/searchcalendar/>

⁵⁵ <https://digi.landesbibliothek.at/viewer/timematrix/>

⁵⁶ Sie werden bei uns meist aus dem Katalog importiert und können im Viewer über den Menüpunkt „Bibliografische Daten“ aufgerufen werden.

⁵⁷ <https://digi.landesbibliothek.at/viewer/metadata/490/1/-/>

⁵⁸ Incunabula Short Title Catalogue. Consortium of European Research Libraries, 2021
https://data.cerl.org/istc/_search

⁵⁹ Gesamtkatalog der Wiegendrucke. Staatsbibliothek zu Berlin – Preußischer Kulturbesitz, 11.08.2021 <https://www.gesamtkatalogderwiegendrucke.de/>

⁶⁰ Ludwig Hain: Repertorium bibliographicum. 1826-1838 <https://digi.landesbibliothek.at/viewer/toc/AC02291478/1/>

⁶¹ https://digi.landesbibliothek.at/viewer!/metadata/Ink-46/1/LOG_0000/

4.2.4. Werkzeuge

In diesem Kapitel werden verschiedene Werkzeuge vorgestellt, die dabei helfen, Digitalisate mit zusätzlichen Daten anzureichern, oder erweiterte Funktionen (z. B. Funktion „Bildausschnitt teilen“) bieten.

4.2.4.1. Bildmanipulation

Die Bildmanipulation kann etwa bei schwer lesbaren Textstellen oder anderen schlecht erkennbaren Darstellungen in alten Drucken und Handschriften eingesetzt werden. Invertieren von Farben, Graustufen, Schärfen, Sättigung, Kontrast, Helligkeit und Färbung kann dabei hilfreich sein, Details (besser) zu erkennen bzw. zuvor vielleicht sogar Verborgenes plötzlich zu entdecken.

4.2.4.2. Bildausschnitt teilen

Um einen Bildausschnitt zu teilen, bieten wir entweder eine Hinweisbox⁶² innerhalb eines Bildes oder den Link eines IIIF-Bildfragments an, welches eher im wissenschaftlichen Bereich zum Einsatz kommen wird⁶³ und in geeigneten Viewern für Vergleichsansichten eingesetzt werden kann.

4.2.5. Merklisten

Merklisten können nur dann benutzt werden, wenn man sich zuvor einmalig registriert hat und angemeldet ist.

Merklisten sind ein Instrument zur Organisation von Werken bzw. auch einzelner Seiten. Diese können über die zugewiesene Merkliste jederzeit rasch aufgerufen werden. Gleichzeitig können diese aber auch von Wissenschaftler:innen zur Darstellung IIIF-fähiger Vergleichsansichten eingesetzt werden.

4.2.6. Mitarbeit

Der Begriff Mitarbeit subsumiert in Goobi den allgemein bekannten Begriff Crowdsourcing. Die momentan eingesetzte Version wurde seit ihrer Erstentwicklung schon mehrmals überarbeitet und visuell an technische Vorgaben und Wünsche der Kund:innen angepasst. Crowdsourcing hat innerhalb der DLOÖ eine große Bedeutung, wurden bisher doch beinahe 70.000 Seiten von externen Benutzer:innen

⁶² Bildausschnitt Pegasus aus Astronomicum Caesareum VD16 A 3074 https://digi.landesbibliothek.at/viewer!/image/AC01107590/17/LOG_0024/#xywh=2500,4873,1253,1237

⁶³ IIIF-Bildfragment Hercules aus Astronomicum Caesareum VD16 A 3074 <https://digi.landesbibliothek.at/viewer/api/v1/re-cords/AC01107590/files/images/00000017.tif/3792,5951,1189,1118/max/0/default.jpg>

bearbeitet. Der Großteil davon wird im Modus „Volltexte bearbeiten“ geleistet. Sollten bei einem Werk auf einer Seite noch keine Volltexte vorhanden sein, dann wird auch der Modus Transkribieren angeboten, bei dem keine Wortkoordinaten⁶⁴ erfasst werden. Zusätzlich stehen noch die Optionen „Inhalte erfassen“ und Kommentare zur Auswahl⁶⁵.

Die neueste Entwicklung im Bereich Crowdsourcing nennt sich Kampagnen. Diese sind im Gegensatz zum herkömmlichen Crowdsourcing vielfältig steuerbar. Mit ihnen lässt sich eine Auswahl bestimmter Werke festlegen und es kann eine zeitliche Befristung eingestellt werden. Für Beiträge kann ein zwingender Review und eine Zugriffslizenz festgelegt werden. Über einen Log können Nachrichten innerhalb einer Kampagne gespeichert und optional allen Teilnehmer:innen zur Verfügung gestellt werden. Eine Statistik gibt einen Überblick darüber, welcher Prozentanteil der Kampagne schon abgearbeitet wurde.

Grundsätzlich kann zum Thema Mitarbeit angemerkt werden, dass unsere bisherigen Erfahrungen äußerst positiv sind. Mitarbeiter:innen sind durchgehend hoch motiviert, eine qualitativ einwandfreie Arbeit abzuliefern, und weit davon entfernt, Arbeit minderer Qualität zu leisten oder vielleicht sogar bewusst falsche Angaben einzuarbeiten. In Zukunft wird sich unser Crowdsourcing verstärkt in Richtung Kampagnen weiterentwickeln, weil diese Form wesentlich granularer justierbar ist.

4.2.7. Kartendarstellungen

Goobi verfügt über ein Karten-Widget, welches OpenStreetMap nutzt und bei Vorhandensein von Geodaten in einem Werk automatisch links der Bildanzeige eingeblendet wird. Geodaten können auf viele verschiedene Arten innerhalb von Goobi erfasst werden.

4.2.7.1. Koordinatenrechteck

Mit Hilfe der Boundingbox⁶⁶ können in Goobi über die Bibliografische Aufnahme Koordinaten aus dem Katalogfeld 034 übernommen werden. Da bei unseren Beständen die meisten Karten als Strukturelemente erfasst werden, hat diese Form der Datenübernahme eher untergeordnete Bedeutung für uns. Die Daten lassen

⁶⁴ Diese bilden die Basis des Highlightings bei Suchanfragen.

⁶⁵ Diese werden unterhalb eines Bildes erfasst und angezeigt <https://digi.landesbibliothek.at/viewer/image/171/1/>

⁶⁶ <https://boundingbox.klokantech.com/>

sich aber auch per Copy and Paste in die betreffenden Koordinatenfelder des Strukturdatums Karte eintragen. Dazu wird zuerst ein Rechteck um das gewünschte Gebiet in der Boundingbox aufgezogen. Die Boundingbox liefert nun die erforderlichen W-O-N-S-Koordinaten.

Die Kartenausschnitte werden, wie in folgenden beiden Beispielen⁶⁷, in einem Widget dargestellt.

4.2.7.2. Verortungen

Für Orte und Plätze können ebenso Geodaten erfasst werden. Orte werden mit einem eigenen Symbol angezeigt. Beide Typen (Orte und Kartenausschnitte) können nebeneinander im Kartenwidget dargestellt werden.

Die Geodaten können über das Feld Geographisches Schlagwort und das Auswahlfeld Geonames⁶⁸ verknüpft werden. Dazu wird der persistente Identifier für einen Ort in das Feld Geonames in Goobi eingefügt⁶⁹. Wird anstatt Geonames im Auswahlfeld GND ausgewählt, so werden die GND-Daten automatisch mit dem Widget verknüpft.

Optional kann auch das Feld Koordinaten verwendet werden. Dieses ist sehr flexibel und benötigt bloß Latitude (Breite) und Longitude (Länge) in folgender Form: Latitude / Longitude. Die Koordinaten kann man sich mit unterschiedlichsten Tools wie z. B. Geonames, GoogleMaps⁷⁰ oder Geoplaner⁷¹ beschaffen.

4.2.8. Bildbereiche

Bildbereiche ist ein neues Feature in Goobi-Production, mit dessen Hilfe man auf einer Seite viele einzelne Strukturen erstellen kann. So lassen sich auf einer Seite mit mehreren Strukturelementen mehrere Abbildungen und Karten als eigene Elemente erfassen. Im Viewer wird dann beim Anklicken eines bestimmten Bildbereiches der Rest des Bildes abgedunkelt. Dadurch wird der Zusammenhang zwischen Text und Bildbereich den Nutzer:innen kenntlich gemacht.

67 https://digi.landesbibliothek.at/viewer/image/AC04601669/34/LOG_0012/ ; https://digi.landesbibliothek.at/viewer/image/AC05371191/105/LOG_0083/

68 <https://www.geonames.org/>

69 Z. B. Gmunden <https://www.geonames.org/2778436>

70 <https://www.google.com/maps>

71 <https://www.geoplaner.de/>

Folgendes Beispiel⁷² zeigt als Hauptstrukturelement eine Karte, auf deren Seite sich mehrere Abbildungen befinden. Diese wurden zuvor in Goobi als Unterstrukturelemente der Karte erfasst. Erkennbar sind diese an dünnen farbigen Umrandungen, sobald man sich auf der Karte befindet. Um nun im Viewer eines dieser Unterstrukturelemente gezielt aufzurufen, kann man entweder direkt im Bild auf einen der farbigen Rahmen oder links unter Inhalt auf die gewünschte Abbildung ([Abb.]: Royan) klicken.

4.2.9. Suche

Suchtreffer werden immer im Kontext des jeweiligen Werkes dargestellt und unterhalb des Werktitels aufgelistet. Die Treffer werden in Metadaten- und Volltexttreffer unterteilt und durch Highlighting im Werk hervorgehoben. Durch diese Form der Organisation können mehrere Suchtreffer innerhalb eines Werkes übersichtlich dargestellt werden.

Die Erweiterte Suche kann aber noch viel mehr. Es können hochkomplexe Suchanfragen mit mehreren Suchgruppen erstellt werden, die ihrerseits wiederum mit Booleschen Operatoren verknüpft werden. Kommt es bei einer Suchanfrage zu einer zu großen Treffermenge, kann diese über unterschiedliche Facettierungen⁷³ eingegrenzt werden.

4.2.10. Kalendersuche

Die Kalendersuche⁷⁴ nutzt das Datumsfeld in Goobi. Das Datumsfeld kann optional bis auf den Tag genau ausgefüllt werden⁷⁵. Sehr häufig wird diese Form der Suche bei unseren Verlustlisten Österreich-Ungarns eingesetzt, wenn die Volltextsuche zu keinem Ergebnis geführt hat, der Zeitraum des Todes aber ungefähr bekannt ist. Im ersten Schritt wird ein Kalenderjahr ausgewählt. Im zweiten Schritt wird ein Jahreskalender angezeigt, bei dem alle Tage mit veröffentlichten Werken direkt zum betreffenden Werk verlinkt sind. So lässt sich rasch eine Verlustliste eines bestimmten Tages aufrufen. Und weil wir jedes Kapitel (Mannschaft, Offiziere, ...) inklusive aller Unterkapitel (nach Alphabet des Familiennamens) erfasst haben, kann man rasch zu einem bestimmten Namen wechseln.

⁷² https://digi.landesbibliothek.at/viewer/image/AC10154567/91/LOG_0043

⁷³ Zeitraum, Sammlung, Dokumenttyp, Person, Erscheinungsjahr, ...

⁷⁴ <https://digi.landesbibliothek.at/viewer/searchcalendar/>

⁷⁵ Verlustlisten Österreich-Ungarns <https://digi.landesbibliothek.at/viewer/search/-/1/CURRENT-NOSORT/DC%3Aperiodika.verlustliste/>

4.2.11. Zeitleiste

Die Zeitleiste⁷⁶ ermöglicht das Browsen durch die Jahrhunderte. Via Schieberegler kann eine Jahresauswahl getroffen werden. Die einzelnen Werke werden durch Anklicken aufgerufen. In der derzeitigen Konfiguration werden maximal 36 Treffer dargestellt.

4.2.12. TagClouds

Sie spiegeln durch ihre Größe die Dichte der jeweils in den Werken vorkommenden Daten wider. Diese werden nach Titel, Orte und Jahre sortiert. Ein Klick auf einen bestimmten Tag öffnet die dargestellte Treffermenge⁷⁷.

4.2.13. Wissenschaftliche Beschreibungen

Für Wissenschaftler:innen bieten wir ein Werkzeug zur Beschreibung einzelner Werke an. Dazu kann ein PDF in den Workflow mithilfe eines vorgefertigten Templates eingespielt werden. Das PDF wird nach dem Hochladen innerhalb der Website in HTML dargestellt. Die Formatierungen des PDFs werden bei diesem Prozess vollständig übernommen, sodass das Layout des PDFs erhalten bleibt. Neben der URN des Werkes wird auch eine PURL⁷⁸ für die Seite angegeben. Sie dient dazu, auf die Seite referenzieren zu können, was für die Zitierbarkeit und wissenschaftliches Arbeiten von Bedeutung ist.⁷⁹

⁷⁶ <https://digi.landesbibliothek.at/viewer/timematrix/>

⁷⁷ <https://digi.landesbibliothek.at/viewer/tags/>

⁷⁸ Das ist eine Möglichkeit, um – wie bei URNs – konstante bzw. persistente Inhalte von Webseiten bereitzustellen.

⁷⁹ <https://digi.landesbibliothek.at/viewer/overview/422/1/>,
<https://digi.landesbibliothek.at/viewer/overview/438/1/>,
<https://digi.landesbibliothek.at/viewer/overview/20/1/>

5. Fazit

Ein Repatorium und die darin enthaltenen Visualisierungen sind immer genau so gut, wie sie betreut, gewartet und weiterentwickelt werden. Um dies zu gewährleisten, sind einerseits personelle und andererseits finanzielle Mittel erforderlich. Repotorienmanager:innen sind dazu aufgerufen, sich intensiv mit der Materie und ihrem System zu beschäftigen. Es genügt nicht, Abteilungsleiter:in zu sein.

Es ist weder notwendig noch zwingend erforderlich, über Programmierkenntnisse zu verfügen oder einen Informatikabschluss zu besitzen, aber es ist unabdingbar, sich mit seinem Repatorium und dessen Möglichkeiten eingehend auseinanderzusetzen. Diejenigen, die ihre Digitalisierungssoftware genau kennen, werden imstande sein, die richtigen Fragen an Fachexpert:innen zu stellen. Nur so wird man erforderliche Entwicklungen erkennen und die richtigen Entscheidungen für die Zukunft treffen.

Bibliografie

- Ciula, Arianna; Eide, Øyvind; Marras Christina; Sahle, Patrick (2018): Models and Modelling between Digital and Humanities. Remarks from a Multidisciplinary Perspective. In: Historical Social Research 43 (4), pp. 343-361.
- Keller, Stefan Andreas; Schneider, René; Volk, Benno (2014): Die Digitalisierung des philosophischen Zettelkastens. In: Keller, Stefan Andreas; Schneider, René; Volk, Benno (Hg.): Wissensorganisation und -repräsentation mit digitalen Technologien. Berlin: de Gruyter Saur, S. 1-17.
- Kogler, Gerald (2014): Barrierefreier Zugang zu offenen Geodaten unter besonderer Berücksichtigung sehbeeinträchtigter Personen. Bachelorarbeit aus Wirtschaftsinformatik. Johannes Kepler Universität Linz. https://www.academia.edu/18954553/Barrierefreier_Zugang_zu_offenen_Geodaten_unter_besonderer_Ber%C3%BCcksichtigung_sehbeintr%C3%A4chtigter_Personen (abgerufen am 10.08.2021)
- Paik, Woojin; Liddy, Elizabeth D.; Yu, Edmund; McKenna, Mary (1993): Categorizing and Standardizing Proper Nouns for Efficient Information Retrieval. In: Boguraev, Branimir; Pustejovsky, James (eds.): Acquisition of Lexical Knowledge from Text. Ohio: State University, pp. 154-160. <https://aclanthology.org/W93-0114.pdf> (abgerufen am 03.02.2023)
- Pamperl, Beate (2017): Visualisierungen in den Digital Humanities – Ein Überblick. In: Maske und Kothurn 63 (1), S. 90-98. <https://dx.doi.org/10.7767/muk-2017-630114>
- Roth, Jeannette (2002): Der Stand der Kunst in der Eigennamen-Erkennung. Mit einem Fokus auf Produktenamen-Erkennung. Lizentiatsarbeit. Universität Zürich, Philosophische Fakultät. <https://www.cl.uzh.ch/dam/jcr:00000000-6a77-a254-0000-00006a27b71e/lizjeannetteroth.pdf> (abgerufen am 16.04.2023)

Windhager, Florian (2017): Choreographien der Existenz. Zur multimodalen Erweiterung biographischer Forschung und Lehre durch Verfahren der visuellen Analyse und Synthese. In: BIOS – Zeitschrift für Biographieforschung, Oral History und Lebensverlaufsanalysen 30 (1-2), S. 60-75. <https://doi.org/10.3224/bios.v30i1-2.06>

Gregor Neuböck leitete die Digitalen Services an der Oberösterreichischen Landesbibliothek bis Oktober 2023. Seit November 2023 leitet er die Stabsstelle Digitales Sammlungsmanagement der Wienbibliothek im Rathaus. Er veröffentlichte unzählige Fachartikel zum Thema Digitalisierung. Zuletzt erschien von ihm in den VÖB-Mitteilungen: 2019 Crowdsourcing an der Oberösterreichischen Landesbibliothek und 2018 sein Buch bei Walter de Gruyter: Digitalisierung in Bibliotheken: viel mehr als nur Bücher scannen.