Treffen des VuFind-Netzwerks der AG FID, 20.09.2023

Protokoll von Anna Lingnau und Patricia Blume

Teilnehmer*innen

- Gunther Kunze (FID Lateinamerika, Karibik und Latino Studies - LACARinfo, Ibero-Amerikanisches Institut)
- Annette Oberreuter (FID Kunstgeschichte - arthistoricum.net, Altertumswissenschaften - Propylaeum, UB Heidelberg)
- Ansgar Teichgräber (FID Altertumswissenschaften - Propylaeum, BSB München)
- Julia Beck (FID Darstellende Kunst, UB Frankfurt am Main)
- Martin Faßnacht (FIDs Tübingen [Theologie, Religionswissenschaft, Kriminologie], UB Tübingen)
- Mario Trojan (FIDs Tübingen, UB Tübingen)
- Angela Pohl (FID Recht)
- Philipp Weiß, FID Altertumswissenschaften - Propylaeum, BSB München
- Tamara Pianos, ZBW, Fachportal EconBiz (nur bis 10:55 dabei)
- Wiebke Herr (FID Geschichtswissenschaft, BSB München)
- Patricia Blume (FID adlr.link, UB Leipzig)
- Bernhard Lutz (FID Musikwissenschaft, BSB München)
- Sophie Straube (FID Osteuropa, BSB München)
- Alexander Purr (FID adlr.link - Kommunikations-, Medien- und Filmwissenschaft, UB Leipzig)
- Nicole Terne (FID Asien, SBB Berlin)
- Arben Hajra (ZBW, EconBiz)
- Dorothée Grieb (FID Altertumswissenschaften - Propylaeum, BSB München)
- Peter Heiss (ULB Münster, Vertretung FID Benelux)
- Kai Ermeling (FID Benelux)
- Konrad Lauenroth (FID Pharmazie, Braunschweig)
- Jannis Ohms (FID Pharmazie, Braunschweig)
- Vladimir Neumann (FID Slawistik, SBB Berlin)
- Maysoun Modzel (FID BBI)
- Anna Lingnau (FID BBI)
- André Lahmann (UB Leipzig)
- Pascal Kanter (FID BBI)
- Olaf Brandt (UB Tübingen)
- Ilona Riek (FID Benelux)

Begrüßung und Organisatorisches

Anna Lingnau begrüßt die Anwesenden zum zweiten FID-VuFind-Netzwerktreffen im Jahr 2023 im Namen der netzwerkinitiierenden Fachinformationsdienste BBI und adlr.link. Sie geht darauf ein,

dass Patricia Blume, die sich kurz vorstellt, im FID adlr.link die Nachfolge des ausgeschiedenen Sebastian Stoppe als Co-Sprecherin übernommen hat. Zugleich weist sie darauf hin, dass sich die beiden Sprecherinnen Unterstützung in Form eines:r Stellvertreters:in wünschen und bittet um diesbezügliche Signale, gerne auch im Nachgang der Veranstaltung.

Im Vorfeld der Sitzung schlug Wiebke Herr (FID Geschichte) vor, die Verwendung der Open Source Lizenz GPL v2 in das Grundlagenpapier des Netzwerkes aufzunehmen. Diese sei ohnehin obligatorisch, da die Lizenz bei VuFind zum Einsatz komme und eine Weitergabe unter gleichen Bedingungen erforderlich sei (Julia Beck). Nachdem Mario Trojan darauf hinweist, dass für Tübinger VuFind-Entwicklungen häufig auch die AGPL3-Lizenz (GNU Affero General Public License) verwendet wird, wird beschlossen, das Thema bei der nächsten Sitzung noch einmal aufzugreifen und intensiver zu diskutieren. Anna Lingnau ruft die übrigen FID dazu auf, die bei ihnen verwendeten Lizenzen (sofern abweichend zu GPLv2) im Vorfeld zu melden.

Vorstellung Personenseiten, Normdatensuche und Self-Archiving Service (Mario Trojan, Tübingen)

Die Präsentation ist einsehbar unter: https://wikis.sub.uni-hamburg.de/webis/index.php/FID-Netzwerk_Vufind

Mario Trojan, Entwickler an der UB Tübingen und zuständig für die FID Kriminologie, Religionswissenschaft und Theologie, führt in das Modul Personenseiten und Self-Archiving Service ein. Die Personenseiten können aus dem Katalog heraus aufgerufen werden und bieten Nutzer*innen die Möglichkeit, Personen, Projekte, Publikationen, Affiliationen und Netzwerke übersichtlich und schnell zu erfassen. Das VuFind-Modul Personenseiten speist sich aus GND-Daten mit Schwerpunkt auf Personen, Körperschaften und Tagungen. Dabei wurde die authority.ini angepasst und um Felder erweitert. Außerdem werden Bilder aus Wikidata importiert und ergänzt. Zudem werden mit verschiedenen Visualisierungen (Tag Cloud, Diagramm) Publikationsleistungen thematisch und auf einer Zeitachse dargestellt.

Beispiel: https://www.ixtheo.de/Authority/134702158

Der in Tübingen entwickelte Self-Archiving Service dockt an den Zweitveröffentlichungsservice an, bei dem der FID personenbezogene Bibliografien (Katalogisierungsprinzip ad personam) erstellt. Auf Basis der bereits im K10plus erfassten Titel können Nutzer:innen nach Authentifizierung die Einträge zu ihrer Person mit Uploads ergänzen, womit ein Admin-Konzept auf Nutzer:innenebene einhergeht. Nach der Abgabe einer rechtlichen Einverständniserklärung und dem Upload in das IxTheo Repositorium erfassen die Bibliothekar:innen die persistenten Identifier (handle und DOI) im K10plus, wo die gesamte Datenhaltung der Tübinger FID passiert. Genutzt wurde DSpace 6, da sich der Release von DSpace 7 verzögerte und mit verschiedenen technischen Schwierigkeiten einhergegangen wäre.

In der Diskussion im Anschluss dankt Anna Lingnau für den Input und vergegenwärtigt den enormen Entwicklungsaufwand in Tübingen. Patricia Blume betont die Vorteile dieses Moduls, weist auf das Interesse des FID-VuFind-Netzwerks sowie speziell die Pläne zu einer Nachnutzung in Leipzig (FID adlr.link, FID BBI) hin. Martin Faßnacht (Tübinger FID) führt in Beantwortung einer Frage aus, dass der Self-Archiving Service noch nicht unter den FID-Nutzer:innen beworben wurde, weil derzeit die Rückmeldung der DFG zum dafür konzipierten Arbeitspaket im Antrag aussteht. Die Community, insbesondere die Wissenschaftsorganisationen, seien sehr interessiert. Julia Beck (FID DK) ergänzt, dass auch der FID DK Normdaten entitätenbasiert verarbeitet, und zwar indem der komplette GND-Datensatz gespiegelt wird. Mario Trojan ergänzt, dass für DSpace 7 bereits erste Anpassungen vorliegen, die interessierten Anwender:innen zur Verfügung gestellt werden können.

Ranking in VuFind und spezifische Herausforderungen für FID (Julia Beck, FID DK)

Die Präsentation ist einsehbar unter: https://wikis.sub.uni-hamburg.de/webis/index.php/FID-Netzwerk_Vufind

Julia Beck (FID DK) gibt eine Einführung in die Funktionsweise des Rankings in VuFind und verweist auf zusätzliche Informationsangebote. Darüber hinaus stellt sie ein von Thomas Wagener und Dennis Schrittenlocher für den HEBIS-Verbund entwickeltes Ranking-Visualisierungstool vor, über das Nutzende die Ursachen für die Positionierung eines Titels im Ranking nachvollziehen können und das auf der WOLFCon vorgestellt wurde. Einsehbar sind Statistiken darüber, wie ein Suchbegriff in den verschiedenen Metadatenfeldern gewichtet wurde und ein Balkendiagramm über die Relevanz eines Titels im Verhältnis zum erstgerankten Titel.

Im Anschluss an die Präsentation skizziert Anna Lingnau Erfahrungswerte und Problemstellungen, die sich beim Ranking im Katalog des FID BBI ergeben: Metadaten, die in Umfang und Komplexität hinter MARC-Daten aus den Verbundsystemen zurückstehen (weil sie zum Beispiel keine Schlagworte oder Klassifikationen beinhalten), werden grundsätzlich niedriger gerankt, unabhängig von der wahrgenommenen inhaltlichen Relevanz für den jeweiligen Suchbegriff. Zusätzlich ist zu beobachten, dass das Ranking Werke mit kurzen Titeln (i.e. Handbücher, Lexika, Überblickswerke) tendenziell "bevorzugt", weil die Relevanz eines Suchbegriffes innerhalb eines Werktitels in Verhältnis zu dessen Länge gewichtet wird, weswegen längere Aufsatztitel in der Regel niedriger gerankt werden. Eine Verbesserung konnte dadurch erreicht werden, dass für Titel jüngeren Erscheinungsjahres eine Boost-Funktion implementiert wurde und die Abstracts in das Relevanzranking mit aufgenommen wurden. Erprobt werden soll auch eine Funktion, über die die zuletzt indexierten Records höher gerankt werden.

Aus der Diskussion im Anschluss:

Auf die Frage, ob jemand bei Usability Tests oder Umfragen in der Zielgruppe auf Mängel des Rankings hingewiesen wurde, antwortet Frau Lingnau, dass User:innen Mängel des Rankings ex negativo mutmaßlich nicht erkennen, diese aber dennoch den Rechercheerfolg beinflussen könnten. Aus Projektkoordinator:innensicht, sei eindeutig festzustellen, dass Records aus bestimmten Datenquellen seltener in der Suche auftauchen als andere. Man könne jedoch erwägen, das Visualisierungstool im Katalog für kurze Zeit im Katalog einsehbar zu machen und das Feedback der Nutzenden hierzu abzuwarten.

Julia Beck weist darauf hin, dass es möglich sein müsste, auch bestimmte Datenquellen im Ranking zu boosten.

Der Pull Request für das Visualisierungstool liegt vor, ist aber noch nicht geprüft: https://github.com/vufind-org/vufind/pull/3069

Update-Runde: Aktuelle Projekte, geplante Neuerungen, Themenvorschläge

Peter Heiss und Kai Ermeling (FID Benelux) beschäftigen sich derzeit mit der Transformation von Metadaten und erwägen, hierfür Benthos zu verwenden. Auf die Frage nach den Alternativen werden die Folgenden vorgeschlagen: Luigi, Metafacture sowie die FID-spezifische Entwicklung Nightwatch (FID Pollux)

Anna Lingnau verweist zusätzlich auf den zu erwartenden Kartierungsbericht des Technik-Boards, in dem die in Hinblick auf bestimmte Themen (Metadatenmanagement, Discovery, Authentifizierungsmechanismen etc.) am häufigsten benutzten Tools skizziert werden.

Vladimir Neumann (FID Slavistik) stellt den Katalog derzeit auf VuFind um und bittet um Austausch in Hinblick auf den Import von Metadaten in VuFind (hier: Dublin Core/XML). Julia Beck verweist auf durch VuFind bereitgestellte Transformationsskripte für verschiedenste Formate, die nachgenutzt werden könnten. Für konkrete Probleme und Fragestellungen könne ein Aufruf im Slack-Channel (http://vufind.slack.com/ , Aufnahme über eine Mail an Chris Hallberg (crhallberg@gmail.com)) des Netzwerkes lohnend sein. Olaf Brandt bietet die UB Tübingen als Ansprechpartner für die Zoteroverfahren an.

Mario Trojan skizziert VuFind-bezogene Vorhaben des FID-Antrags, der der DFG derzeit zur Begutachtung vorliegt. Die Präsentation ist einsehbar unter: https://wikis.sub.uni-hamburg.de/webis/index.php/FID-Netzwerk_Vufind

- So soll das Result Grouping der UB Leipzig nachgenutzt und auf die neueren VuFind-Versionen angepasst werden. André Lahmann (UB Leipzig) bietet hier seine Unterstützung an.
- Einbindung der im ConfIDent-Portal der TIB Hannover eingetragenen wissenschaftlichen Veranstaltungen in das FID-Portal.
- Einbindung von Zitationsdaten von OpenAlex in den FID. Hier verweist Jannis Ohms auf die entsprechende Entwicklung des FID Pharmazie und bietet Unterstützung an.
- Inhaltliche Relevanzfilterung (bedingt VuFind-spezifisch): Es stellt sich die Frage, welche Untermenge der Nachweise einer externen Datenbank zu unseren aktuellen Nachweisen passt? Am Beispiel von KrimDok und CORE möchten die Tübinger um eine Ähnlichkeitsanalyse von Fremd- und Bestandsdaten bemühen und freuen sich über Beteiligungen anderer FID.
- Automatisierte Auszeichnung von Metadaten (bedingt Vufind-spezifisch): Geplant sind eine automatisierte Vergabe von Klassifikationen in Metadaten sowie ein Mapping von Klassifikationen via Cocoda. Hier verweist Anna Lingnau auf ein ähnliches (nicht bewilligtes) Vorhaben des FID BBI und meldet Interesse an einem gemeinsamen Vorgehen.

Anna Lingnau und Alexander Purr berichten aus den aktuellen Projekten der VuFind-AG der UB Leipzig: Zum einen sei derzeit geplant, die schon bei EVIFA implementierte Fernleihfunktion in den eigenen Katalog zu integrieren und hierbei eine direkte Weiterleitung des Nutzenden zum passenden Verbundkatalog zu ermöglichen (ohne dass der Nutzende selbst den Verbund auswählen muss). Hierfür werden Daten (Heimatbibliothek der Nutzenden) aus der für VuFind entwickelten Benutzerverwaltung fidis nachgenutzt.

Zum anderen arbeitet der FID adlr.link an einem Visualisierungstool für Suchergebnisse, dass voraussichtlich in der nächsten Sitzung vorgestellt werden könne.

Julia Beck empfiehlt die aufschlussreiche Präsentation der UB Leipzig (André Lahmann) auf der WOLFcon zu „Security Insights From Analyzing VuFind Code“ (https://www.youtube.com/watch?v=x2UbTFPk7-8).