Protokoll zum 2. Arbeitstreffen des FID-Netzwerks Philologien

11. Mai 2023, 9:00-11:00 Uhr, virtuell
https://etherpad.studiumdigitale.uni-frankfurt.de/p/Netzwerk_Philologien

Organisation:
Michel, Volker, FID AVL, v.michel@ub.uni-frankfurt.de
Mrotzek, Arne, FID AVL, a.mrotzek@ub.uni-frankfurt.de
Renner-Westermann, Heike, FID Linguistik, h.renner-westermann@ub.uni-frankfurt.de
Dimitrova, Vanya, FID Linguistik, v.dimitrova@ub.uni-frankfurt.de

Teilnehmer*innen:
Adam, Volker, FID Nahost, volker.adam@bibliothek.uni-halle.de
Adamczak, Katarzyna, FID Ost, katarzyna.adamczak@bsb-muenchen.de
Bauer, Axel, FID Nahost, axel.bauer@bibliothek.uni-halle.de
Eggert, Eric, FID Philosophie, eggert@ub.uni-koeln.de
Einert, Katharina, FID Lateinamerika, Karibik und Latino Studies, fid@iai.spk-berlin.de
Fritze, Thorsten, FID Linguistik, t.fritze@ub.uni-frankfurt.de
Geißler, Nils, FID Philosophie, nils.geissler@uni-koeln.de
Grüter, Doris, FID Romanistik (Doris.Grueter@ulb.uni-bonn.de)
Hampf, Yannik, FID Philosophie, hampf@ub.uni-koeln.de
Hass, Sophia, FID Karten, fid-karten@sbb.spk-berlin.de
Imeri, Sabine, FID Sozial- und Kulturanthropologie, sabine.imeri.1@ub.hu-berlin.de
Kaplan, Jonathan, FID Jüdische Studien, j.kaplan@ub.uni-frankfurt.de
Neumann, Vladimir, FID Slawistik, vladimir.neumann@sbb.spk-berlin.de
Paterson, Duncan, FID Asien, duncan.paterson@sbb.spk-berlin.de
Riek, Ilona, FID Benelux, riek@uni-muenster.de
Schuller, Dorothea, FID Anglo-American Culture, schuller@sub.uni-goettingen.de
Schumann Douosson, Anne, FID Afrikastudien, a.schumann@ub.uni-frankfurt.de
Sell, Leonore, FID Anglo-American Culture, sell@sub.uni-goettingen.de
Trapp, Markus, FID Romanistik, markus.trapp@sub.uni-hamburg.de
von der Krone, Kerstin, FID Jüdische Studien, k.vonderkrone@ub.uni-frankfurt.de
Weiß, Philipp, FID Altertumswissenschaften - Propylaeum, philipp.weiss@bsb-muenchen.de

Gäste:
Glaab-Kühn, Frederike, Kompetenzzentrum für Lizenzierung (SBB), friederike.glaab-kuehn@sbb.spk-berlin.de
Koenig, Pauline, Kompetenzzentrum für Lizenzierung (SBB)
Piel, Pattrick, Max Weber Stiftung/OPERAS-GER, piel@maxweberstiftung.de
Töpfer, Marlen, Max Weber Stiftung/OPERAS-GER, Toepfer@MaxWeberStiftung.de
Voigtschild, Fabian, Kompetenzzentrum für Lizenzierung (BSB), fabian.voigtschild@bsb-muenchen.de

TOP 1 Begrüßung & Protokoll
Die Sitzung wird durch Volker Michel (FID AVL) eröffnet. Er begrüßt im Namen der Organisator*innen alle Teilnehmer*innen und stellt kurz die Tagungsordnung vor. Das Protokoll des letzten Netzwerk-Treffens wird verabschiedet und auf webis veröffentlicht (vgl. https://wikis.sub.uni-hamburg.de/webis/images/b/bd/Protokoll_FID-Netzwerk_Philologien_2022-11-17.pdf).

TOP 2 Vorstellung der Europäischen Forschungsinfrastruktur OPERAS
Referent*innen: Marlen Töpfer und Pattrick Piel (OPERAS-GER)
Link zur Präsentation: https://wikis.sub.uni-hamburg.de/webis/images/9/9f/OPERAS-GER_FID_Philologien_2023_05_11.pdf

OPERAS ist eine europäische Infrastruktur zur Stärkung offener Wissenschaftskommunikation in den Geistes- und Gesellschaftswissenschaften. OPERAS ist ein Konsortium, das über 50 Mitglieder umfasst. ORERAS-GER fungiert seit 2020 als nationale Kontaktstelle (National Node) zur Vernetzung der europäischen Forschungsinfrastruktur in der deutschen Wissenschaftslandschaft. Das Projekt wird vom BMBF gefördert und ist an der Max-Weber-Stiftung angesiedelt (vgl. https://operas-ger.hypotheses.org/).
Die Referent*innen präsentieren die Governance Structure des Konsortiums sowie die Special Interest Groups und die Möglichkeiten, in die Infrastruktur hineinzuwirken. Die angebotenen Services und die Möglichkeiten zur Nachnutzung werden vorgestellt. Ein besonderer Fokus liegt dabei auf der multilingualen Discovery-Plattform GoTriple (vgl. https://www.gotriple.eu). Die Suchfunktionalitäten werden anhand von konkreten Beispielen live demonstriert.

Fragen und Diskussion
Renner-Westermann: Werden nur OA-Publikationen indexiert?
Piel: Es werden nur OA-Dokumente nachgewiesen.

Renner-Westermann: Das Vokabular, das für die Suche verwendet wird, umfasst circa 3000 Begriffe. Wie wurden diese ausgesucht? Gibt es Konkordanzen für die Zuordnung der Dokumente?
Töpfer: Es handelt sich um eine intellektuelle Auswahl, bei der alle Mitglieder-Institutionen involviert waren (vgl. https://www.gotriple.eu/ontology/triple, https://www.semantics.gr/authorities/vocabularies/SSH-LCSH/?language=en)
Piel (im Chat): Wenn man das Vokabular ansieht, sind auch die entsprechenden Sprachen angegeben. GoTriple umfasst Croatian, Dutch (partial), English, French, Finnish, German, Greek, Italian, Polish, Portuguese, Spanish, Ukrainian (vgl. https://www.gotriple.eu/about)

Renner-Westermann: Wie viele Datenquellen bzw. Dokumente sind über GoTriple verfügbar?
Töpfer: über 9.4 Mio. Dokumente, 22.000 Projekte, 8 Mio. Profile. BASE wird auch geharvestet (aktuell ca. 4 Mio. BASE-Treffer; die Integration ist noch nicht abgeschlossen).

Adam (im Chat): Könnten auch Digitalisate mit hinterlegtem, OCR-erkanntem Volltext in GoTripel integriert werden?
Töpfer: Aktuell nicht, aber grundsätzlich möglich; theoretisch können alle Dokumente, die textbasiert sind, hinterlegt werden.

Renner-Westermann: Ist eine Mitgliedschaft im Konsortium Voraussetzung für die Kooperation zwischen OPERAS und FID? Welche Kosten sind damit verbunden?
Töpfer: Für die Kooperation und Nutzung der Services ist keine Mitgliedschaft notwendig; grundsätzlich ist Mitgliedschaft willkommen (die Mitgliedschaft-Modelle variieren, z. B. jährliche Gebühren oder Selbstverpflichtung, Personal zur Verfügung zu stellen).

Mrotzek: Wer kann sich an den Special Interest Groups beteiligen?
Töpfer: Diese Gruppen sind offen; Kontakt aufnehmen und mitmachen (vgl. https://operas-eu.org/special-interest-groups/)

Schuller: In der Trefferliste werden englische Titel angezeigt, auch wenn die Originaltitel Deutsch sind. Werden die Titel automatisch übersetzt? Ist das eine Voreinstellung, die man ausschalten kann?
Töpfer: Aktuell wird alles übersetzt und es ist nicht möglich, in der Trefferliste direkt die Originalsprache anzuzeigen (diese wird erst in der Detailansicht sichtbar); dieser Nutzer-Wunsch wird an die Entwickler weitergegeben.

Mrotzek: Verschiedene Länder scheinen in OPERAS unterschiedlich stark aktiv zu sein. Wie kommt es dazu?
Töpfer: Ausschlaggebend ist die Struktur des Konsortiums: Core Member tragen mehr Verantwortung und beteiligen sich intensiver an den Entwicklungen. Das Konsortium wird von Frankreich koordiniert, wo auch die Idee für OPERAS entstanden ist.

von der Krone: Unter den angebotenen Sprachen fehlen zentrale osteuropäische Sprachen; hier ist eine Erweiterung wünschenswert.
Welches Datenmodell wird verwendet? Europeana Data Model (EDM) oder ein anderes?
Piel: GoTriple hat ein eigenes Datenmodell; Information darüber findet sich unter https://zenodo.org/record/7359654#.ZFx93HZBw2w

von der Krone: Mit welchen Identifiern werden die Forscherprofile verknüpft? Sind sie mit ORCID verlinkt?
Töpfer: Über eine Verknüpfung mit ORCID wird diskutiert, ist aber noch nicht implementiert.

Mrotzek: Was wäre das Ziel eines OPERAS-Workshops mit den FID?
Töpfer: Es wäre interessant von den FID zu hören, was für sie im europäischen Forschungsraum wichtig ist; ein gemeinsamer Workshop wäre dafür gut geeignet.
Mrotzek: Dies wäre nicht nur für die im Netzwerk Philologien vertretenen FID interessant, sondern für einen breiteren Kreis; die Idee sollte in der großen FID-Runde vorgestellt werden.

Michel: Das Projekt OPERAS-GER endet voraussichtlich im September 2023. Was passiert nach dem Projektabschluss?
Töpfer: Die weitere Förderung ist noch ungewiss. Auch wenn OPERAS-GER nicht mehr als Projekt gefördert wird, bleibt die Max-Weber-Stiftung selbst als Core Member von OPERAS weiterhin als Kontaktstelle in Deutschland bestehen.

TOP 3 Metadaten - (direkter) Austausch, Export, Nachnutzung
Referent: Vladimir Neumann (FID Slawistik)
Link zu Präsentation: https://wikis.sub.uni-hamburg.de/webis/images/d/d2/2023_FID-AG-Philologien-Metadaten_Neumann.pdf

Zentralisierter Austausch via KfL und direkter Metadatentausch zwischen FID werden einander gegenübergestellt.
Zuerst stellt Neumann schematisch den KfL-Workflow dar:
Das KfL fungiert als Sammelstelle für Metadaten von FID- und Nationallizenzen. Wenn vertraglich vereinbart, liefern die Produkt-Provider Metadaten ans KfL (grobe Produktbeschreibung, Zugangslinks). Die Daten werden vom KfL gesichtet und evaluiert, bei ausreichender Qualität nach MARC konvertiert und in die GBV-Datenbanken (FIDELIO) eingespielt. Es gibt Filter-Mechanismen; der gefilterte Content kann in das FID-eigene System integriert werden.
Es stellt sich die Frage, ob ein direkter Metadatentausch notwendig ist. Argumente dafür:

- Metadaten liegen oft in fremden Sprachen vor; die sprachliche Expertise, die für eine hohe Qualität der Konvertierung der Daten gebraucht wird, ist in den FID vorhanden.
- Diese hochspeziellen Daten erfordern auch spezialisierte Behandlung; z. B. Unicode-Handling, Diakritika (dafür spezielle Indexierung), diverse Transliteration- und Transkriptionsformate.
- Die Sacherschließungssysteme können zuverlässig von den FID-Experten gemappt werden.
- Feinere Filter-Mechanismen können implementiert werden (z. B. um aus Datenbanken allgemeiner Natur nur philologische Inhalte zu selektieren).
- Konversionsdaten haben eine komplexere Struktur als normale bibliografische Daten

Fazit: Der direkte Metadatentausch zwischen FID ist schneller als über den GBV; die Selektion erfolgt mit schärferem Profil.

Das Team im Slavistik-Portal hat 18 Jahre Erfahrung im Bereich Datenhandling und Datenaufbereitung. Es findet z. B. eine Anreicherung um die ZDB-ID statt – dadurch ergeben sich weitere Möglichkeiten und Funktionalitäten.
Neumann präsentiert auch Überlegungen zur Frage, wie ein Metadatenaustausch innerhalb des Netzwerks in der Praxis funktionieren könnte. Wünschenswert wäre eine Netzwerk-interne Übersicht, welche Metadaten die einzelnen FID im Einsatz haben (Vorschlag: auf der Netzwerk-Seite können die einzelnen FID ihre Übersichten verlinken, dadurch kann eine Art Metadaten-Börse entstehen).
Der FID Slawistik braucht für die Konvertierung 1-2 Wochen. Die aufbereiteten Daten werden dann auf einem FTP-Server zum Download bereitgestellt. Diese Form des Austauschs (Benennung des Contents, des gewünschten Zielformats und der Filterkriterien) funktioniert gut in der Praxis; Erfahrung mit FID Ost.

<u>Fragen und Diskussion</u>
Adam: Viele Parallelen, was die Bedarfe der FID betrifft. Problematisch ist allerdings, dass es lange dauert, bis die lizenzierten Produkte in den anderen Verbünden nachgewiesen werden (Nicht-FID-Bibliotheken beschweren sich).
Neumann: Wenn Lizenzen ablaufen, dann sind die Daten nicht mehr verfügbar und müss(t)en aus den Verbünden gelöscht werden. Für die Verbünde ist das ein großer Aufwand. Deswegen besser auf FIDELIO setzen oder lokale Konvertierung leisten und in die Suche integrieren.
Glaab-Kühn: Daten liegen in K10+ und sind über verbundübergreifenden Datenaustausch einzubinden; dies funktioniert allerdings zurzeit nicht gut, lieber direkt aus FIDELIO. (FIDELIO ist eine eigene Kollektion bei K10plus-Zentral).

Renner-Westermann: Wie geht man mit Lizenzbedingungen / fachlich definierten Nutzerkreisen um, wenn man die FID-Lizenzen eines anderen FID im eigenen Portal nachweisen will?
Glaab-Kühn: Der Nutzerkreis ist im Vorfeld vertraglich festgelegt; Nachweis mit Vermerk auf die Nutzerkreisdefinition des anderen FID; Verlinkung zu DBIS vorhanden, und in DBIS findet sich ein Link zur Definition des Nutzerkreises; grundsätzlich muss die Registrierung immer über den lizenzierenden FID laufen.
Neumann: Damit man direkt über die entsprechenden HAN- oder Proxysysteme zugreifen kann, muss man die Zugangslinks lokal anpassen (ist das überregional realisierbar?)

Paterson: Sind die im Rahmen des Slavistik-Portals erstellten Skripte und Mappings irgendwo publiziert?
Neumann: Die Skripte sind zwar auf GitHub, aber noch nicht öffentlich zugänglich; Kontakt mit Neumann aufnehmen – er stellt sie zur Verfügung.

Renner-Westermann: Haben Sie Personalkapazitäten, um Export-Wünsche der FID bedienen zu können?
Neumann: Die Werkzeuge sind entwickelt. Bei klar definierten Selektionskriterien ist der Aufwand überschaubar.

Grüter: Es ist problematisch, dass Datennachweise nicht in die anderen Verbünde (oder nur verzögert) kommen; FIDELIO in KVK einbinden als Lösung?
Glaab-Kühn: KVK lehnte eine Einbindung ab, weil die Nutzerkreise fachlich definiert sind; bei Interesse seitens der FID kann das KfL in der neuen Projektphase einen zweiten Versuch unternehmen. – Das KfL wird Mitte 2023 eine FID Info „Metadaten und FID“ publizieren.

Adam (im Chat): KVK-Einbindung von FIDELIO wäre ein sinnvoller Schritt, unsere Lizenzen sichtbarer zu machen.
von der Krone (im Chat): Für Forschende ist der KVK ein zentrales Rechercheinstrument, oft wichtiger als die Verbundkataloge.

Neumann: Sowohl zentralisierter Austausch als auch lokaler Austausch sind sinnvoll; am besten in beide Richtungen – bereinigte Daten können dann in die Verbünde zurückgespielt werden.
von der Krone: FIDELIO-Qualität reicht nicht aus; Vorschlag: FIDELIO als Lösung für „quick and dirty“, später Bereinigung und Anreicherung der Daten durch die FID und Zurückspielen an FIDELIO.

TOP 4 Integration Roundtable FID/Text+ sowie Bericht zum Frühjahrstreffen von Text+
Referent: Eric Eggert (FID Philosophie)
Am Anfang stand die Frage, wo man Materialien und Protokolle des Roundtable hosten könnte. Das Netzwerk Philologien bietet sich hier wunderbar an, da es auch eine große Überschneidung der verpartnerten Institutionen gibt. Die Netzwerk-Treffen eignen sich darüber hinaus gut für thematische Vor- und Nachbearbeitung der Roundtables und für die Abstimmung im Netzwerk. Künftig wird die Dokumentation zum Roundtable auf der webis-Seite des Netzwerks Philologien veröffentlicht. Die Protokolle der vergangenen Sitzungen wurden bereits publiziert.

Im März fand ein internes Treffen der Arbeitsgruppen und Task Areas (TA) des Konsortiums Text+ statt (27.-28.03.2023, Leipzig). Insgesamt war ein größeres Bewusstsein für das Verhältnis zwischen FID und NFDI zu beobachten, jedoch unterschiedlich ausgeprägt in den verschiedenen TA von Text+. Es gilt weiterhin, Kommunikationsprozesse zu intensivieren. Es gibt drei FID-Vertreter*innen in den Gremien von Text+:

- Dr. Anna Lingnau (FID Buch-, Bibliotheks- und Informationswissenschaft) / Scientific Coordination Committee Collections
- Eric Eggert (FID Philosophie) / Operations Coordination Committee
- Heike Renner-Westermann (FID Linguistik) / Operations Coordination Committee

In der Sitzung der AG FID Koop wurde der Aktionsplan des FID Netzwerks (des FID LG) vorgestellt. Es wurde thematisiert, welche Bedeutung der Plan für Text+ hat und wie das Konsortium damit umgehen soll. Das Prinzip der synergetischen Komplementarität wurde dabei mehrfach betont. Anknüpfungspunkte und gemeinsame Ziele wurden definiert, z. B.:

- Vernetzung der technischen Infrastruktur als Voraussetzung für die Verstetigung

- Vermittlerfunktion der FID zwischen Fachcommunity und NFDI
- LOD als gemeinsames Feld

Was kann man gemeinsam machen (neben Gremienarbeit):
Cordi (Conference on Research Data Infrastructure, 12.-14.09.2023, Karlsruhe, vgl. https://www.nfdi.de/cordi-2023/); Poster zu Verhältnis zwischen FID und NFDI (ein Vorbereitungsgespräch findet am 15.05.2023 statt; Interessierte aus dem Kreis können gerne mitwirken); wichtig, dass die FID an der Tagung teilnehmen und Präsenz zeigen.
Geplant ist auch eine Kartierung der FDM-Dienste der verpartnerten FID (weitere Informationen folgen).
Am 17.05.2023 findet der zweite FID/Text+ Jour Fixe zum Thema Consulting statt (vgl. https://events.gwdg.de/event/378/)
Am 23.06.2023 findet am Leibniz-Institut für Deutsche Sprache in Mannheim ein Text+-Workshop zum Thema „Wohin damit? Storing and reusing my language data“ statt (vgl. https://events.gwdg.de/event/372/); FID sind eingeladen.

TOP 5 Sonstiges

Michel: Die Umfrage "Ein Service in einem Satz" fand wenig Resonanz, was vermutlich am ungünstigen Zeitpunkt kurz vor Antragseinreichung lag; bei Kartierungsinitiativen soll künftig auf das Community-Wiki verwiesen werden; zudem ist es immer möglich, besondere Services im Rahmen der Netzwerk-Treffen vorzustellen.

Michel informiert über Angebote des open-access.network: Im Rahmen des BMBF-geförderten Projekts open-access.network (s. https://open-access.network/ueber-uns/oa-network) planen das Helmholtz Open Science Office der Helmholtz-Gemeinschaft, die Technische Informationsbibliothek und das Open-Access-Büro Berlin, angesiedelt an der UB der Freien Universität Berlin, in Zusammenarbeit mit wissenschaftlichen Fachgesellschaften Workshops und Informationsgespräche zur Stärkung von Open Access beim wissenschaftlichen Publizieren anzubieten.
Interesse an gemeinsamen Veranstaltungen kann bis 30.06.2023 bei Linda Martin (linda.martin@open-access-berlin.de) gemeldet werden.

Mrotzek: Auf der XIX. Tagung der Deutschen Gesellschaft für Allgemeine und Vergleichende Literaturwissenschaft (30.05. – 02.06.2023, Potsdam, vgl. https://zwischenspiele.de/) veranstaltet der FID AVL einen Workshop „Zwischenstände. Neue Literatur- und Informationssysteme für die Philologien“ (in Kooperation mit Text+, FID Philosophie, FID Romanistik, FID Musikwissenschaften, FID Darstellende Kunst, FID Lateinamerika, Karibik und Latino Studies, FID Anglo-American Culture, FID Medien-, Kommunikation- und Filmwissenschaft).
Für die Workshop-Teilnehmer*innen werden Goodie bags vorbereitet. Falls ein FID Werbematerialien beisteuern möchte, gerne den FID AVL kontaktieren.

Das nächste Netzwerk-Treffen findet im Herbst 2023 statt (bevorzugt nach der Buchmesse). Eine Terminumfrage folgt.