

## Protokoll der 12. Sitzung der FAG EI

am 10. Dezember 2020 als Online-Sitzung, 13.00 - 15.00 Uhr

### Anwesende

Uma Balakrishnan  
Regine Beckmann  
Barbara Block (ab 13:35 Uhr)  
Anke Boehrnsen  
Heike Carstensen  
Reiner Diedrichs  
Berrit Genat  
Silke Horny  
Michaela Probst  
Kai Steffen

### Entschuldigt

Sabine Bruck  
Bettina Kunz

### Protokollführung

Anke Boehrnsen

### Tagesordnung

**Top 1:** Verabschiedung des Protokolls der FAG-EI-Sitzung vom 26.10.2020

**Top 2:** Wiki der FAG EI

**Top 3:** Berichte von den Online-Tagungen.  
Welche Themen resultieren daraus für die FAG EI?

- ["4. Workshop Computerunterstützte Inhaltserschließung"](#)
- [„Erfahrungen und Perspektiven mit dem Toolkit Annif“](#)

**Top 4:** Kurze Berichte aus den Gremien

- RDA-Entwicklung im D-A-CH-Raum (Horny)
- AG K10plus Katalogisierung (Horny)
  - [Veröffentlichte Protokolle](#)
- AG K10plus Sacherschließung (Balakrishnan)
  - [Protokoll AG KXP SE 20201123.docx](#)
- Fachbeirat (Beckmann)
- VZG / Ziel- und Leistungsvereinbarung 2021 (Diedrichs)

**Top 5:** Abschluss oder Übergabe der Arbeitspakete in die nächste Amtszeit

## Top 6: Verschiedenes

- Aktueller Stand des coli-conc-Projektes (Balakrishnan)
- RVK im GVK (Balakrishnan, Steffen)

## Protokoll:

### Top 1: Verabschiedung des Protokolls der FAG-EI-Sitzung vom 26.10.2020

### Top 2: Wiki der FAG EI

Die neue Struktur des Wikis der FAG EI wird vorgestellt und dankbar angenommen.

### Top 3: Berichte von den Online Tagungen:

- **"4. Workshop Computerunterstützte Inhaltserschließung"**
  - Aus dem über K10plus hinausgewachsenen Anwenderkreis des DA-3 und benachbarten Projekten wurden spezifische Nutzungsszenarien präsentiert, u.a.
    - die Aufsatzerschließung und Ansätze auf dem Gebiet multilingualer Verfahren im FID Theologie
    - die Erschließungsmethoden des Fachinformationsbundes FIV als neuem Mitglied mit eigenem Vertrag
    - die Anpassungen für die Anforderungen der ZBW
    - die aktuellen Entwicklungen in der DNB (Culturegraph, Erschließungsmaschine EMA).
  - Der DA-3 wird sich weiterentwickeln zu einer zentralen Plattform für die Verarbeitung von Vorschlägen aus ganz unterschiedlichen, auch maschinellen Verfahren. Ein weiterer Schwerpunkt ist geplant für die umfangreiche Einbindung lokaler Systeme.  
Im abschließenden Anwendertreffen wurde neben dem Thema Datenkennzeichnung noch über ein zukünftiges Vermittlungskonzept und über die Weiterentwicklung der RVK-Funktionen im DA-3 beraten.  
Es ist vorgesehen, in O-BIB einen Tagungsbericht zu veröffentlichen.
  - Wichtige Diskussionspunkte des Workshops und Themen, die in der FAG EI aufgegriffen werden könnten, sind:
    - Wie können zukünftig multilinguale Verfahren gestaltet werden? Der FID Tübingen z.B. schreibt Übersetzungen in zahlreichen Sprachen in die GND-Datensätze (bisher nur auf der K10plus Sicht). Problematisch ist aber aus Sicht der Teilnehmer:innen dabei, dass Äquivalenzen aus einfachen Übersetzungen in einem Fachgebiet nicht bedeuten, dass die Begriffe auch in anderen Fachgebieten passen. Gebraucht würde stattdessen die qualifizierte Verknüpfung zu Normvokabularen mit Thesaurusrelationen.
    - Hieran, aber auch an die Integration maschinell erzeugter Daten in das Vorschlagstool, schließt das wichtige Thema einer differenzierten und aussagekräftigen zukünftigen Datenkennzeichnung in K10plus an. Im

Macs Projekt wurde z.B. nicht der Grad der inhaltlichen Übereinstimmung zwischen GND Schlagwort und Übersetzungsterm angegeben. Solche Angaben werden aber sowohl für die Datenerzeugung als auch für die Datenpräsentation zunehmend wichtiger. Ebenso für die Beurteilung und Weiterbearbeitung maschinell erzeugter Vorschläge. (Wo wird die Provenienz der Daten in den Metadaten abgelegt? Können Konfidenzwerte am Deskriptor zugefügt werden? Könnte ein Bewertungssystem in den intellektuellen Vorgang integriert werden?)

- Ein interessantes Projekt sollte beobachtet werden: Martin Braschler präsentierte ein Evaluationsprojekt der Zürcher Hochschule für Angewandte Wissenschaften (ZHAW) zusammen mit Eurospider zur Wirkung von Inhaltserschließungsdaten auf das Retrieval. Dabei werden Daten unterschiedlicher Herstellungsverfahren untersucht. Bisher scheint es aber noch keine aussagekräftigen Ergebnisse zum Zusammenspiel der Qualität von Inhaltserschließung und dem Retrievalresultat zu geben.
- **„Erfahrungen und Perspektiven mit dem Toolkit Annif“**
  - Zweitägige Veranstaltung im Rahmen des "[Netzwerks maschinelle Verfahren in der Erschließung](#)"
  - Tag 1: [Einführungsvortrag](#) über automatische Unterstützungstools für automatische und inhaltliche Erschließung (Koninklijke Bibliotheek, Den Haag), Werkstattberichte aus [TIB](#), [ZBW](#) und [DNB](#) zur Evaluierung von Annif, 2 Diskussionsgruppen: "[Einsatzmöglichkeiten von Annif](#)" und "Weiterentwicklung von Annif"
  - Tag 2: [Vorstellung](#) des Toolkit (entsprach dem [Tutorial](#) von Osma Suominen und Mona Lehtinen für die FAG EI am 03.03.2020 ), Vorstellung der [Weiterentwicklung](#) im Rahmen eines Forschungsprojekts in der ZBW, Impulsvortrag "[Gemeinsamer Werkzeugkasten](#)", Diskussionsgruppen "DACH-Kooperation maschinelle Verfahren" und "Infrastruktur zur maschinellen Erschließung mit Annif"
  - Die Werkstattberichte haben praxisnah dargelegt, wie man mit dem Toolkit umgehen kann und welche Verknüpfungsmöglichkeiten zu vorhandenen Anwendungen bestehen. Aus dem Bericht der TIB geht hervor, dass die intellektuelle und technische Vorarbeit erheblich mehr Aufwand erfordert, als der tatsächliche Einsatz von Annif. Quintessenz: Die Daten müssen eine bestimmte Beschaffenheit haben, um mit Annif arbeiten zu können. Trainingsdaten und Vokabular müssen zunächst umfänglich und inhaltlich definiert und beschafft werden, wobei das Exportformat eine entscheidende Rolle spielt. Für die Erstellung von Vokabular (an dem trainiert wird) und Trainingsdatenset müssen die Daten aufbereitet werden, da Annif bestimmte Formate verlangt (Textbasiert oder SKOS). Für die Aufbereitung der Daten sind Vorkenntnisse erforderlich, u.a. in Python (oder GREL), SKOS, RDF-Triple-Erzeugung, Nutzung von "Library Carpentries" wie OpenRefine oder Protégé. Idealerweise könnte dieser Schritt eingespart werden, wenn die Daten von vornherein die Anforderungen für Annif erfüllen: z.B. GND in SKOS, Sprachcodes enthalten, die Langform enthalten zu den (z.B.) BK-Notationen. Alternativ könnten Transformations-Features in Annif integriert werden, dies

wäre noch zu prüfen und ggf. zu entwickeln. DNB, ZBW, TIB werden Annif weiter evaluieren. Die ZBW hat bereits ein Bewertungsverfahren integriert, in der TIB wird evaluiert, ob die automatische Fächerzuordnung mit dem Averbis-Algorithmus durch Annif abgelöst werden kann.

- In den Diskussionen kam immer wieder die Frage auf, welche "Vorleistungen" durch die GND erbracht werden können, z.B. SKOSifizierung der GND, verbindliche Richtlinien für die Erfassung von Normdaten. Die Einsatzmöglichkeiten sind abhängig von den Evaluierungsergebnissen der Institutionen, die Annif bereits einsetzen. Ziel soll sein, den "Werkzeugkasten" Annif niederschwellig auch für Bibliotheken nutzbar zu machen, die nicht über Ressourcen wie TIB, ZBW und DNB verfügen. Eine offizielle "DACH-Kooperation" wird eher nicht angestrebt - es wird zu viel Overhead befürchtet. Die als "Guerilla-Aktionen" bezeichneten Aktivitäten in einzelnen Einrichtungen sollen ggf. durch bilaterale Kooperationen zusammengeführt werden. Angestrebt wird ein gemeinsamer Datenpool für den Austausch bzw. die Bereitstellung von Trainingsdaten und Vokabularen.

Diskussion: Welche Themen resultieren aus den Tagungen für die weitere Arbeit in der FAG EI?

Die FAG EI wird wie bisher Empfehlungen zu einzelnen Tools aussprechen. Die Entwicklung und Evaluierung der Tools wird in der FAG EI beobachtet; in Bezug auf Annif berichten TIB und ZBW direkt in die FAG EI. Hinsichtlich der Struktur und des Inhalts der Metadaten lassen sich aus den Evaluierungsergebnissen konkrete Anforderungen für die maschinelle Nachnutzung in Erschließungstools ableiten (s.o.: Datentypisierung (inkl. Sprachkennzeichnung), Verwendung von Standards, Vollständigkeit). "Negativ"-Beispiele für die Erschwernis einer maschinellen Nachnutzung sollten gesammelt und wenn möglich an geeigneter Stelle publik gemacht werden, um diesbezüglich eine größere Sensibilisierung zu erreichen. Die umfassende inhaltliche Erschließung von Beständen zu ermöglichen bleibt weiterhin das Ziel der FAG EI, wobei Evaluierungsprojekte wie das oben zitierte von Martin Braschler beobachtet werden. Ein verbundübergreifendes valides Erschließungskonzept wäre sinnvoll.

#### **Top 4: Kurze Berichte aus den Gremien**

- **AG K10plus Katalogisierung**

- Die AG K10plus Katalogisierung hält monatliche Telefonkonferenzen ab. Es werden hierbei Fragen aus der K10plus-Praxis geklärt. Entscheidungen werden in die K10plus-Katalogisierungshandbücher eingearbeitet. Die Protokolle der AG werden [veröffentlicht](#).
- In der AG K10plus Katalogisierung wird regelmäßig über das 3R-Projekt berichtet und die praktische Regelwerksarbeit verfolgt. Die FG Erschließung hat den Vorschlag gemacht, ein Erschließungshandbuch zu erarbeiten, in das die bisherigen RDA-Regelungen und DACH-AWR überführt werden. Die Erarbeitung soll kooperativ erfolgen. Das Erschließungshandbuch wird sich in drei Teile gliedern: Allgemeine Anleitungen, Entitätenbeschreibung und Ressourcentypen. Neue Konzepte und Regelungen aus den überarbeiteten RDA werden nur übernommen, wenn sie erkennbar einen Mehrwert für die

Erschließung und den Nutzer bieten und sie einfach in die Bibliothekssysteme implementiert werden können.

- **AG K10plus Sacherschließung**

Die Mitglieder der AG K10plus Sacherschließung treffen sich i.d.R. zweimal bis dreimal im Jahr. Das zweite Treffen des Jahres fand am 23.11.20 online statt. Die Hauptpunkte der Diskussion bei diesem Treffen waren u.a.:

Umgang mit veralteten RVK-Notationen im K10plus und deren Kennzeichnung , Kennzeichnung der mit dem DA erschlossenen Titeldatensätze und Kennzeichnung der angereicherten SE durch Mappings aus coli-conc , Feld für coli-conc Mappings in den TK-Sätzen, Stand der Aktualisierung der BK,, Nachspielung der DNB-SE und Stand der Print-Online Synchronisation. Die ausführlichen Protokolle der Sitzungen sind auf der Seite

<https://opus.k10plus.de/solrsearch/index/search/searchtype/collection/id/16302> veröffentlicht.

- **Fachbeirat**

- Zu jedem Protokoll der FAGs wird ein Management Summary erstellt und direkt bei den Protokollen verlinkt. Die Management Summaries werden durch die VZG an die Bibliotheksleitungen verschickt, um dort die Wahrnehmung der Arbeit der GBV-Gremien, der engen Zusammenarbeit mit der VZG und die Transparenz zu steigern. Im nächsten Jahr werden Aufwand und Bedarf noch einmal evaluiert. Für die praktische Arbeit ist zu berücksichtigen, dass die Management Summaries in der Regel sehr viel schneller benötigt werden, als die Protokolle fertig sind. Doppelarbeit ist die Folge.
- Schwerpunkt der Sitzung war die Ziel- und Leistungsvereinbarung. Die in der letzten Sitzung der FAG EI besprochenen Maßnahmen wurden aufgenommen und die Terminierung ggf. angepasst. Die Entwurfsversion wurde an den FB verteilt.
- Aus der Diskussion:
  - Bei der Synchronisation von Sacherschließungsdaten zwischen Print- und Online-Titeln ist im BSZ eine Teillösung aktiv, die übernommen werden könnte. Synchronisiert werden in größeren Abständen die Titel, indem die PPNs in Feld 4243 der Onlineresourcen ausgewertet werden. Wichtig, aber komplizierter und nicht so zügig machbar, ist weiterhin die Synchronisation von nicht verlinkten Daten.
  - Die Kennzeichnung der OA-Titel muss rückwirkend für die Altdaten durchgeführt werden.
  - Wie können die Altdaten maschinell angereichert werden?

- **VZG / Ziel- und Leistungsvereinbarung 2021**

## **Top 5: Abschluss oder Übergabe der Arbeitspakete in die nächste Amtszeit**

- **Erstentwurf zur Ideensammlung (Steffen):** Die Ideensammlung soll von den FAG-Mitgliedern bis 21. Januar 2021 kritisch geprüft, um weitere Ideen ergänzt und ggf. kommentiert werden. Sie wird in der Sitzung der nächsten Amtszeit diskutiert.

- Dazu soll auch **Verbliebenes** aus den letzten Arbeitspaketen geprüft und für die Arbeitspakete ab 2021 ergänzt, Erledigtes markiert und die interne Seite ins Archiv übertragen werden.
- Wesentliche und grundsätzliche Punkte sollen in die Wikiseite "**Inhaltliche Schwerpunkte 2021-2023**" übernommen werden.

## Top 6: Verschiedenes

### Aktueller Stand des coli-conc-Projektes

Für die abschließende Phase von 2021 bis 2023 hat das coli-conc Projekt eine weitere Förderung von der DFG bewilligt bekommen. Mit den zusätzlichen Personalmitteln beabsichtigt die VZG, die Weiterentwicklung des Projektes voranzutreiben und als nachhaltigen Dienst zu etablieren.

coli-conc hat seit September 2020 eine neue Webseite. Im November wurde der BARTOC Dienst von der VZG übernommen, er wird von coli-conc weitergepflegt. In den Monaten Oktober und November 2020 veranstaltete das Team mehrere Workshops, u.a. im Rahmen der SWIB-Konferenz und WikiCite.

Im nächsten Schritt sind geplant:

- Einbindung von BARTOC in Cocoda
- Einbindung der coli-ana Ergebnisse in Cocoda
- Konvertierung weiterer Haussystematiken und Einspielung in Cocoda
- Die vorhandenen Mappings in die WinIBW integrieren und ihre Nachnutzung erleichtern,
- Automatische Anreicherung der RVK-BK Mappings mittels des Moduls *coli-rich*
- Einbindung von coli-conc an den DA und
- Weiterentwicklung der neuen coli-conc Webseite

- **RVK im GVK:**  
Bei der Verknüpfung der RVK-Notationen mit den Normsätzen, die vorher nur als Texteinträge im Titelfeld 5090 vorlagen, wurde festgestellt, dass diese nicht suchbar waren, um neue Treffermenge zu bilden. Der Fehler lag an dem angewendeten Suchschlüssel IKT 5040, dieser wurde ausgetauscht und der Fehler behoben. Im ersten Quartal 2021 ist geplant, auch die Suchbarkeit der Hierarchien durch deren Verlinkung im Feld 553 zu ermöglichen.
- Sitzungen der FAG EI in 2021  
Aufgrund der Corona-Pandemie werden auch im nächsten Jahr überwiegend virtuelle Sitzungen stattfinden. Wenn es die Umstände zulassen, wird eine Präsenzsitzung abgehalten werden.
- Die GBV-Verbundkonferenz 2021 wird auch virtuell stattfinden (wegen nicht ausreichender Ressourcen für die Tagungsorte)

**Nächster Termin als Online Meeting: 21.1.2021 um 13 Uhr**

**Thema: Konstituierung der FAG EI 2021 - 2023**