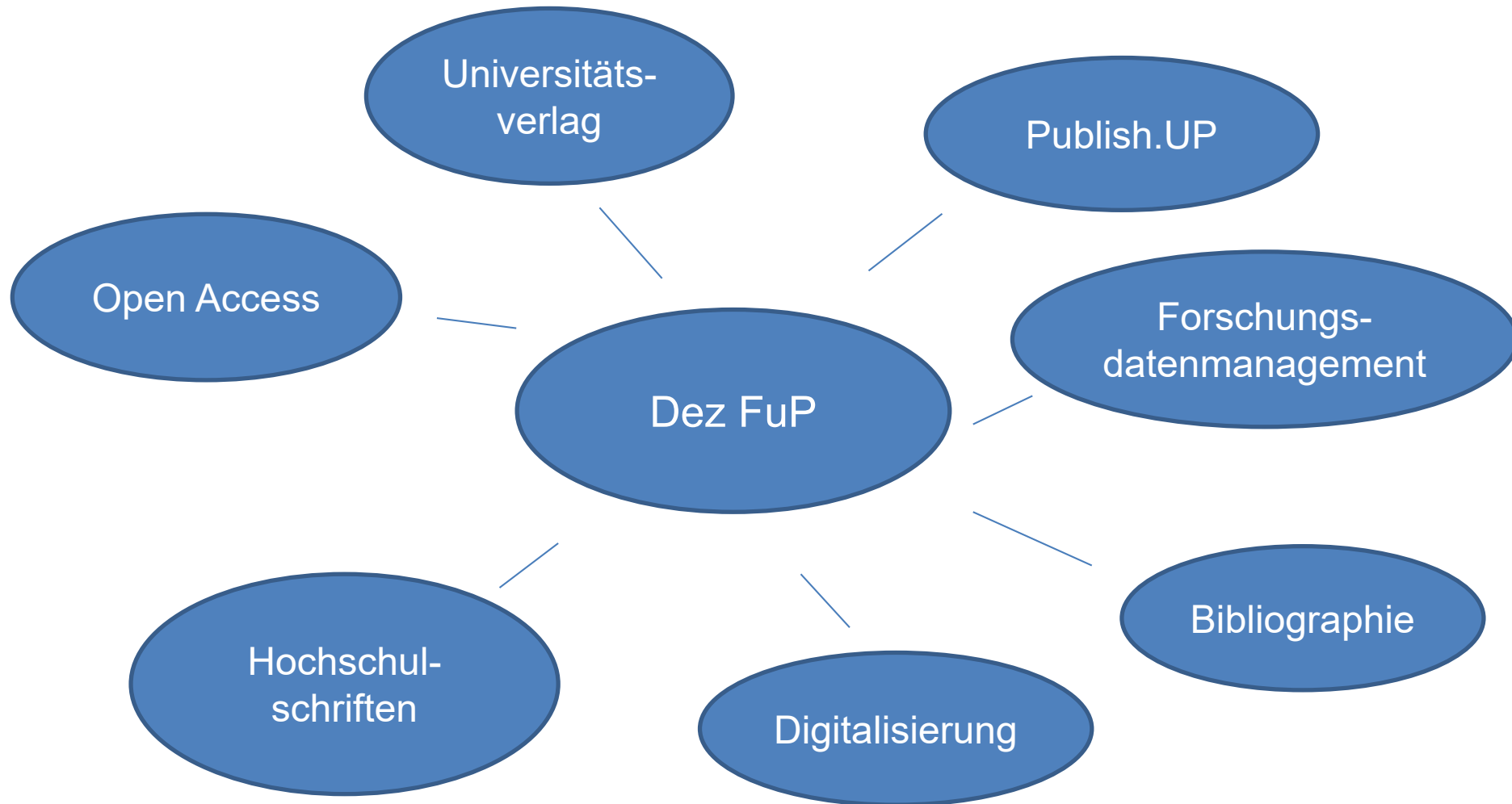


Import bibliographischer Metadaten in Publish.UP

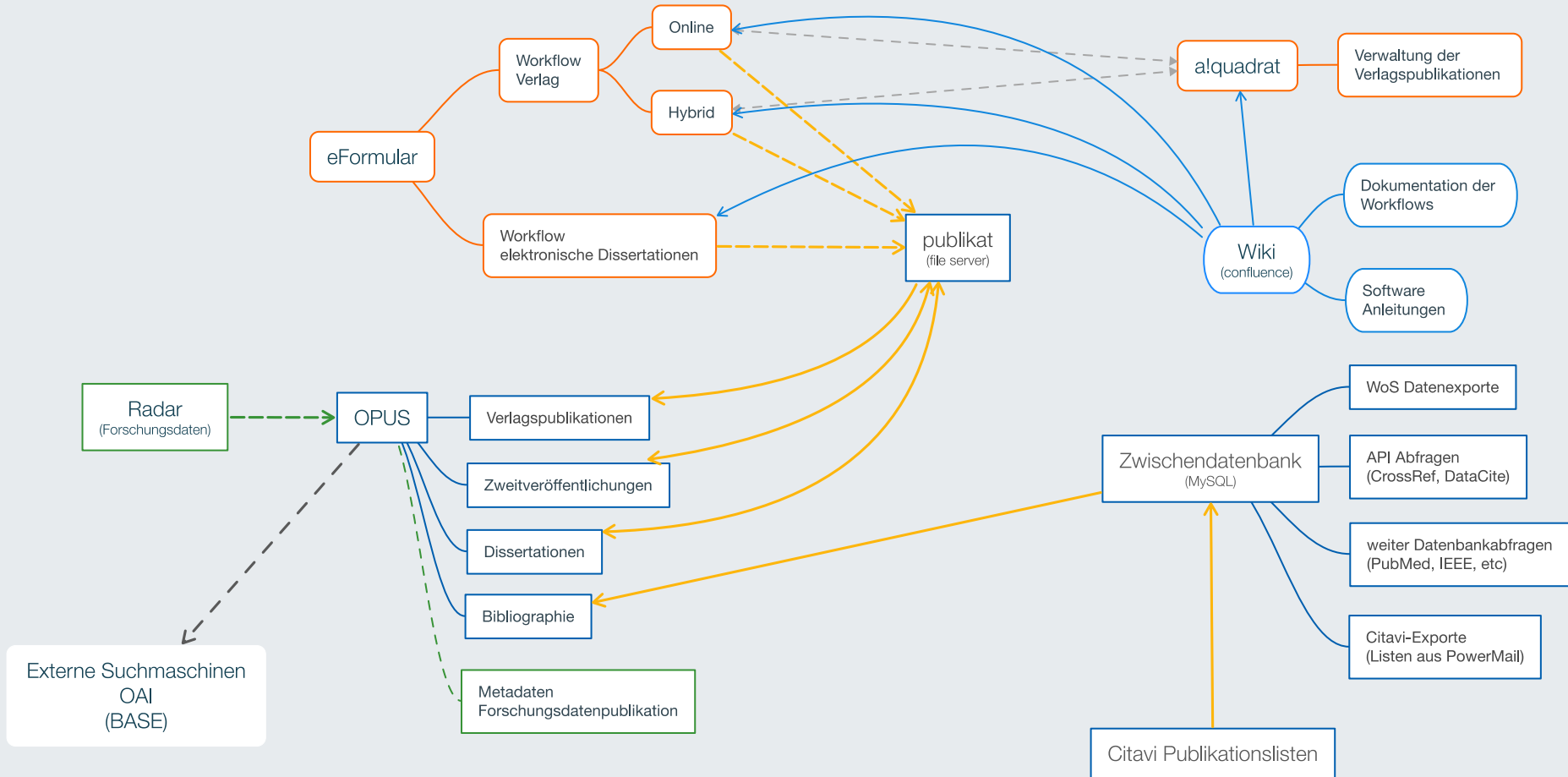
Marco Winkler

OPUS 4-Workshop im Rahmen des
109. Bibliothekartages
am 15.06.2021

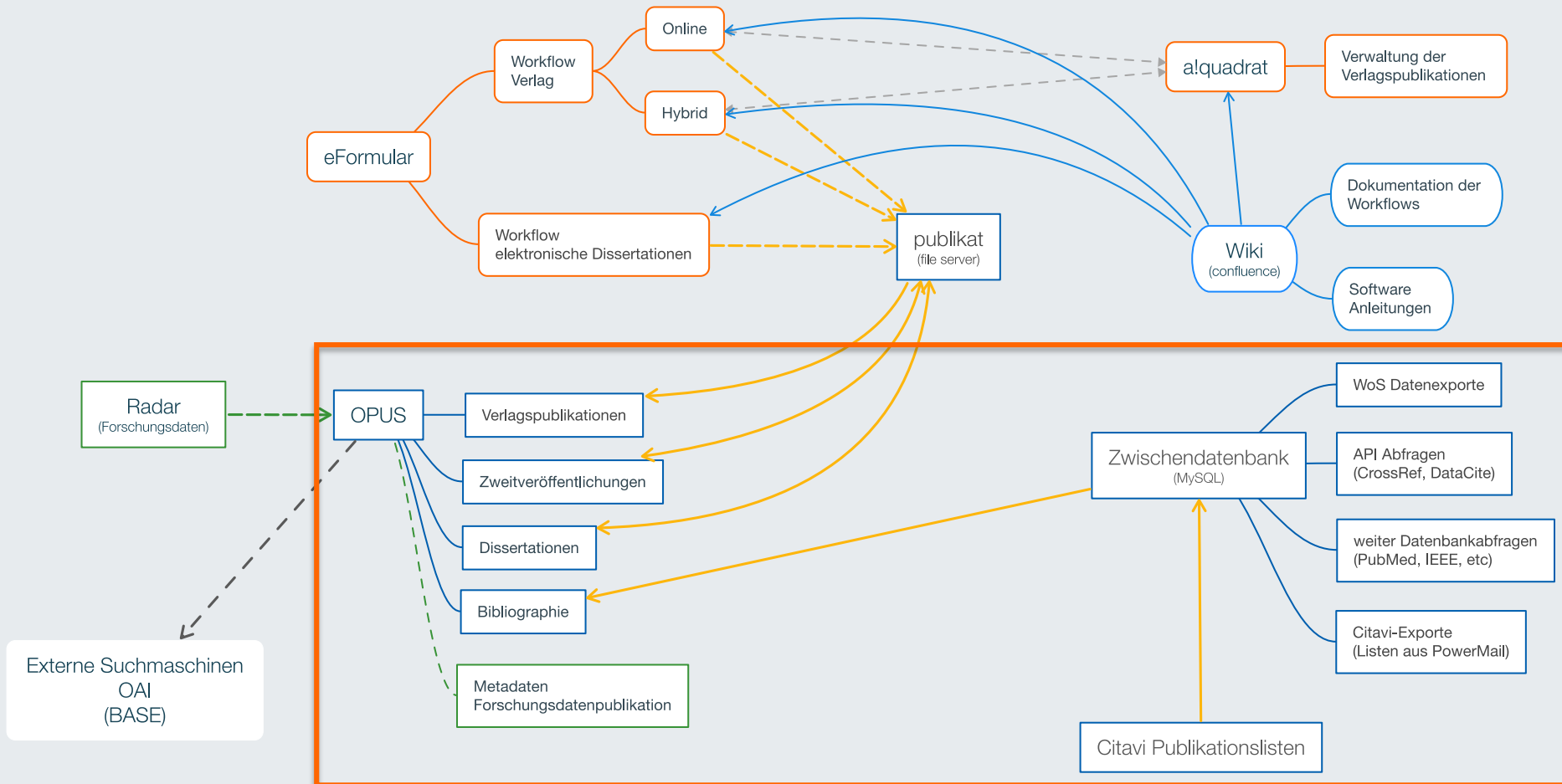
Das Dezernat Forschungs- und Publikationsunterstützung



Datenhaltende Systeme (Bibliographie, Publikationsserver)



Datenhaltende Systeme (Bibliographie, Publikationsserver)

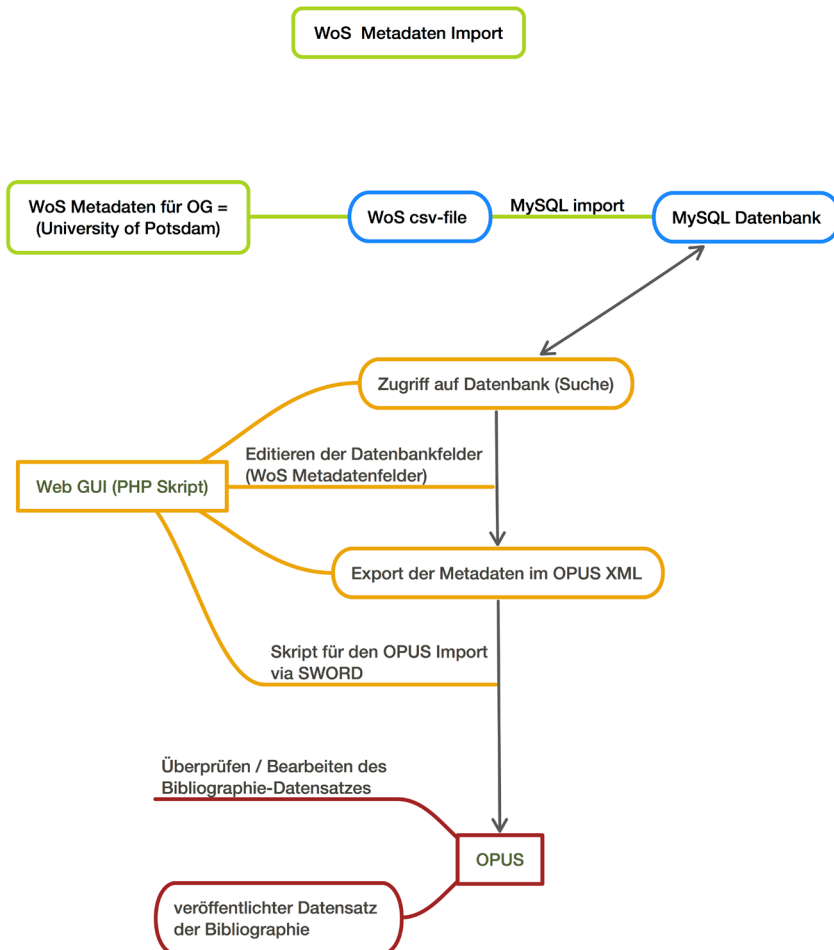


- Publikationsserver für Textpublikationen und die Universitätsbibliographie der Universität Potsdam
- Publikationssystem OPUS in Version 4.7.0.4 gehostet vom KOBV
- Verwaltet ca. 47.800 Datensätze (11.700 Volltexte, 36.100 Bibliographiedatensätze)
- Ca. 85 eingerichtete und dokumentierte Datenfelder zur Beschreibung der Ressourcen

Bibliographieanreicherung und Zweitveröffentlichung

- Import von Publikationsmetadaten über SWORD-Schnittstelle
- Daten aus verschiedenen Quellen: Citavi, WoS, CrossRef, Output.UP (Anbindung geplant, Eigenentwicklung der UBP zur Ermittlung des Publikationsaufkommens der Universität und dessen Finanzierung), Unpaywall, DOAJ API...
- Importsystem zur Anreicherung und redaktionellen Bearbeitung der importierten Daten (MySQL-Datenbank und Weboberfläche)
- Seit 2015: Massenimporte über die Importschnittstelle im Format OPUS XML
- Seit 2019: Einzelimporte über SWORD

Fluss der Publikationsdaten



- Publikationsdaten aus externen Quellen werden in SQL-DB eingespielt
- Web-GUI ermöglicht die Bearbeitung der Metadaten vor Einspielung in OPUS
- Dabei wesentliche Anreicherung mit Angaben zur Instituts- und Fakultätszugehörigkeit

Anreicherungen

Publikationsdaten werden beim Import angereichert und normalisiert, z.B.:

- **Personen:** Name in Ansetzungsform der GND, ORCID, GND-ID, E-Mail-Adresse der Universität, Corresponding Author, akad. Titel
- **Zeitschriftentitel, Verlag:** Ansetzungsform der ZDB
- **Fakultät, Institut:** redaktionelle Zuordnung unter Nutzung einer Historie
- **Open-Access-Status:** automatisch ermittelt und redaktionell geprüft
- **CC-Lizenz:** DOAJ (Journalebene), redaktionell geprüft

Anreicherungen

- **Identifizier:** ISSN, DOI, Pubmed-ID, WoS-Accessionnummer
- **Klassifikation:** DDC, JEL
- **Funding:** Förderer, Förderprogramm
- **Deskriptive Metadaten:** Schlagwörter, Abstracts
- **Strukturelle Metadaten:** Relation zu Postprints
- **Peer Review-Status:** redaktionelle Zuordnung

- **Collections, Enrichments:** Nutzung der Collection-Ids und Enrichmentnamen von OPUS für den Import (Datenformat OPUS XML)
- **Ersetzungstabellen:** für Verlagsnamen, Verlagsorte und Zeitschriftentitel werden Konkordanztabellen genutzt, die heterogene Bezeichnungen unterschiedlicher Datenquellen auf die bevorzugten Ansetzungsformen mappen

Bearbeitung im Importsystem

- Formular zur Bearbeitung der Metadaten im Importsystem mit Ausfüllhilfen, z.B. Abfragen der Unpaywall-API zum OA-Status
- Button „Speichern und Produktivinstanz“ exportiert den Datensatz nach OPUS

Fakultät 68*: Bitte wählen... ▼

Institut 68A*: Bitte wählen... ▼

Zweites Institut 68B: Bitte wählen... ▼

fertig? (0=nein; 1=ja)*:

vorhanden? (0=nein; 1=ja):

Document Type neu:

Institut 68A_coll_old:

Institut 68A_coll_b_old:

OA Marker Import:

OA:

DOAJ-API-Abfrage:

DOAJ-API-Abfrage-unpaywall:

oa-Status-unpaywall:

oa-Status-UP-Kriterien:

cc ISSN-DOAJ-Abfrage:

PM:

[Datensatz ohne Speichern verlassen](#)

Bearbeitung in OPUS

- Normalisierung und Anreicherung in OPUS, z.B.:
 - GND-Ids für Personen
 - Namensansetzungen
 - Corresponding Author
- Bibliographiedatensatz:

<https://publishup.uni-potsdam.de/50903>

Benutzerdefinierte Felder (Enrichments) + Hinzufügen

Schlüssel	Text	
Datenquelle	wos:2020	Entfernen
wos_accessionNumber	WOS:000491142000001	Entfernen
reprint_address	Gago, J (reprint author), UIB, Inst Agroecol & Econ Agua INAGEA, F	Entfernen
Fördernde Institution	Ministerio de Economía y Competitividad (MINECO), Spain [CTM20]	Entfernen
opus.import.date	2021-06-03T07:49:54+00:00	Entfernen
opus.source	sword	Entfernen
opus.import.user	importub	Entfernen
opus.import.file	filename=package.tar	Entfernen
opus.import.checksum	e291760b0bce20a7e3c8a4aae663bc	Entfernen
corresponding_author	Gago, Jorge	Entfernen

Workflow: Anleitung

Checkliste zur Bearbeitung

Aufzeichnung der Schulung vom 16.12.2020

Nr.	Arbeitsschritt	Bearbeitungshinweise
1.	<p>Aufruf des nächsten freien Datensatzes im gewünschten Datenbereich oder Suche nach einer bestimmten Publikation</p>	<p>Das Skript ruft den nächsten Datensatz auf, der nicht durch einen anderen Bearbeiter gesperrt bzw. noch nicht nach OPUS importiert (Feld "transferopus" = 0) oder manuell auf erledigt gesetzt (Feld "fertig" = 1) wurde.</p> <p>Die Suche erfolgt derzeit entweder in den Tabellen 2017-2019, 2006-2008 oder goldoa und ist mit Autorennamen, DOI und Titeln möglich. Titelstichwörter können in der Reihenfolge gesucht werden. Trunkiert wird mit "%".</p>
2.	<p>Publikationsdaten in OPUS suchen:</p> <p>Gibt es bereits einen Bibliographieeintrag für den Titel?</p> <p>a) Wenn ein Bibliographieeintrag vorhanden ist:</p> <p>aa) Enthält der Bibliographieeintrag bereits alle zu erfassenden Metadaten?</p> <p>Dann das Feld fertig? (0=nein; 1=ja)*</p> <p>auf 1 setzen und Button "Nur Speichern" klicken.</p> <p>bb) Fehlen im Bibliographieeintrag auf OPUS Metadaten?</p> <p>Dann fehlende Metadaten ergänzen, insbesondere:</p> <ul style="list-style-type: none"> • wos_accessionNumber • Article Number • Fördermittelgeber • reprint_address • WOS_authorAddress • Corresponding Author • Schlagwörter • Pubmed-Id • ISSN • etc. <p>Alternativ den WoS-Datensatz importierten und den vorhandenen Datensatz löschen.</p> <p>Achtung: Sollte der zu löschende Datensatz bereits ein Enrichmentfeld "Externe Bemerkung" haben, dann muss im neuen Bibliographiedatensatz die Relation neu gesetzt UND im referenzierenden Datensatz die Relation auf den importierten Bibliographiesatz geändert werden (siehe zum Setzen der Relation unter 4.).</p>	<p>Gibt es bereits ein Postprint der Publikation, aber keinen Bibliographieeintrag?</p> <p>Bei Schritt 4. an die Relationen über das Enrichmentfeld "Externe Bemerkung" denken, z.B.:</p> <p>Bibliographiedatensatz:</p> <p>Zweitveröffentlichung in der Schriftenreihe Postprints der Universität Potsdam : Humanwissenschaftliche Reihe ; 613</p> <p>Postprint:</p> <p>Bibliographieeintrag der Originalveröffentlichung/Quelle</p>

Workflow: Erfassungsleitfäden

Wissenschaftlicher Artikel

Erstellt von Kerstin Schröder, zuletzt geändert von Marco Winkler am Jul 13, 2020

Reihenfolge in Administration	Erfassungspriorität (1 2 3 4 5)	OPUS Feld	Label	Datentyp	Kardinalität (0 1 n)	Pflichtfeld in Anmeldeformular	Metadatenstandard	Erfassungsbeispiel
		Allgemeine Metadaten						
1	1	Language	Sprache	VARCHAR(255)	1	1	ISO 639-2	Englisch
2	1	Type	Dokumenttyp	VARCHAR(100)	1	1	Gemeinsames Vokabular für Publikations- und Dokumententypen / DINI, DNB, BSZ http://nbn-resolving.de/urn:nbn:de:kobv:11-100109998	Wissenschaftlicher Artikel
		publication_state	Status	ENUM('draft', 'accepted', 'submitted', 'published', 'updated')	1	0		published
		Zeitbezogene Metadaten						
5	3	CompletedDate	Veröffentlichungsdatum (online)	VARCHAR(50)	1	0	dd.mm.yyyy	06.06.2016
6	1	CompletedYear	Jahr der Fertigstellung	SMALLINT(4)	1	1	yyyy	2017
3	1	PublishedDate	Datum der Erstveröffentlichung	VARCHAR(50)	1	0	dd.mm.yyyy	06.06.2016
4	1	PublishedYear	Jahr der Erstveröffentlichung	SMALLINT(4)	1	0	yyyy	2017
		server_date_created		VARCHAR(50)	1		dd.mm.yyyy	06.06.2016
		server_date_modified		VARCHAR(50)	1		dd.mm.yyyy	06.06.2016

Fazit

- Übernahme von Publikationsdaten für die Bibliographie und Zweitveröffentlichungen aus freien oder lizenzierten Datenquellen vereinfacht die Identifikation relevanter Publikationen und deren Erfassung in OPUS
- Unabdingbar bleibt eine ressourcenaufwändige, redaktionelle Bearbeitung der importierten Daten (Normalisierung, Standardisierung, Anreicherung), um die nötige Datenqualität zu gewährleisten.
- Wir leben nicht im bibliographischen Schlaraffenland: Leider keine „sich selbst füllende Bibliographie“

- Technische Anpassung des Datenmodells von OPUS wünschenswert, z.B.
 - Personenmanagement (Personen nur einmal im System führen und verknüpfen)
 - Entitäten für Verlage, Institute, Klassifikationen etc.
 - Identifier für Institutionen
 - Affiliations: Verknüpfung einer Person mit Institutionen
 - Relationen: Abbildung von Dokumentbeziehungen, z.B. Werkauflagen, Artikel und Zweitveröffentlichung etc.
 - Datentypen zur Validierung und Vermeidung fehlerhafter Einträge

Kontakt

Marco Winkler
Universitätsbibliothek Potsdam
Dezernat Forschungs- und Publikationsunterstützung
Universitätsverlag

E-Mail: marco.winkler@uni-potsdam.de
Tel.: +49 331 977-2032

<https://www.ub.uni-potsdam.de/de/publizieren/uebersicht>