

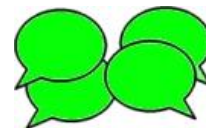
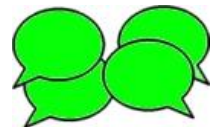
# *Workshop #3*

## *Thema Migration*

Felix Hemme, ZBW | Uschi Klute, VZG | Ingolf Kuss, hbz  
12. April 2019

# Agenda

- „Thinking transformation...”
- Datenmigration
- Metadaten
- (technische) Migrationsverfahren
- (formale) Datentypen
- (thematische) Datengruppen
- User-Daten in FOLIO
- Migration von Nutzerdaten



„Thinking  
transformation...”



*“Thinking transformation  
is more important than  
data migration”*

- Migration = Umstellungsprozess eines Informationssystems
- Prozess → Projekt
- Projektablauf in der Bibliothek
  - Projektorganisation
  - Fachliche Arbeitsgruppen
- FOLIO-Terminologie

# Datenmigration



# Datenmigration - primäre Fragestellungen

In unserem lokalen Bibliothekssystem haben wir ...

- (formale) Datentypen
- (thematische) Datengruppen
  - modulübergreifend
  - modulspezifisch
- Welche Daten wollen wir übernehmen?
  - migrieren? → ETL-Prozess
  - manuell erfassen?



# (formale) Datentypen - Einführung



# Datentypen

## Beschreibende Metadaten

- Titeldaten
- Normdaten
- Bestandsdaten
  - Lokaldaten / Holdings
  - Exemplardaten / Items

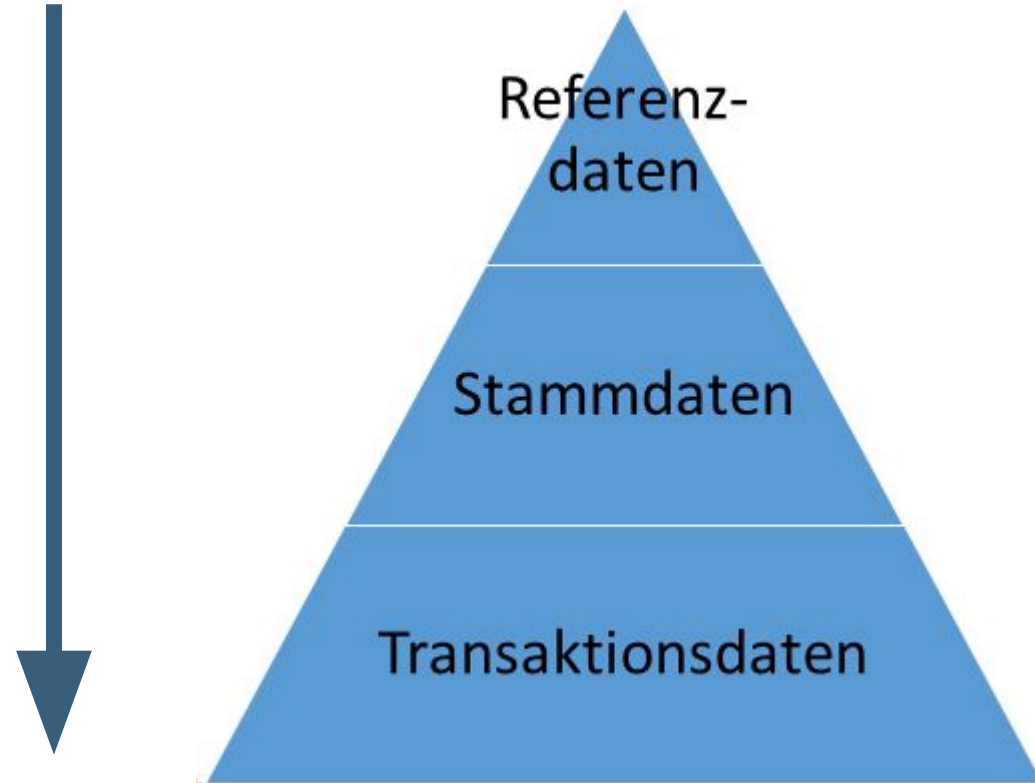
## Daten im lokalen Bibliothekssystem

- Referenzdaten
- Stammdaten
- Transaktionsdaten



# Datentypen im Bibliothekssystem

- Mehr Datensätze
  - Kürzere Lebensdauer
  - Abnehmende Bedeutung der Datenqualität
- Reihenfolge des Datenimports



# Bibliografische Metadaten



# Beschreibende Metadaten aus der Katalogisierung

## Titeldaten

- enthalten Informationen über ein Buch, eine Serie, eine Tonaufnahme, eine Videoaufzeichnung usw.
- Erfassung der Daten gemäß **Regelwerk** in einem **bibliografischen Datenformat** anhand eines **Schemas** und Ausgabe in einer **Kodierung**.<sup>1</sup>
- Beispiel
  - DE: Regelwerk RDA | Pica3/+ | ../pica.dtd | Pica XML
  - US: Regelwerk RDA | MFBD | ../MARC21slim.xsd | MARCXML

<sup>1</sup> Nach Voß et al.: <https://format.gbv.de/about>

# Beschreibende Metadaten aus der Katalogisierung

## Normdaten

- Enthalten standardisierte Formen für Namen, Titel und Schlagwörter, die in bibliographischen Datensätzen verwendet werden
- Beispiel:
  - DE: Regelwerk RDA und RSWK für Schlagwortkatalogisierung, Erfassung in diversen Formaten (Pica3, MAB 2, MFAD), zentrale Normdatei GND
  - USA: Regelwerk RDA, Erfassung hauptsächlich im Format MFAD, zentrale Normdatei “Library of Congress Authorities / NAF”

# Beschreibende Metadaten aus der Katalogisierung

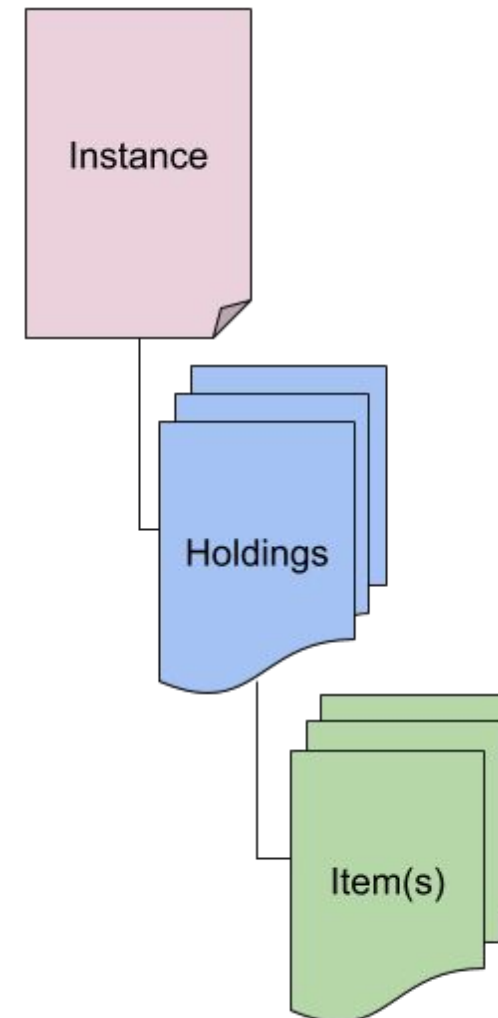
## Bestandsdaten

- weisen den Bestand einer Bibliothek an einer bestimmten Ressource nach und sind abhängig von einem Titeldatensatz
  - DE: unterschiedlich: im GBV keine Unterscheidung zwischen Holdings und Items, Bestandsdaten sind direkt in den bibliografischen Datensatz integriert; ZDB: Nutzung von MFHD als eigenständige Datensätze; hbz: eigene Records für Holdings und Items
  - US: teilweise werden MFHD genutzt, Unterscheidung zwischen Holdings (z.B. eine Signatur) und Items (alle Exemplare zu dieser Signatur), mehrere Items sind somit mit einem Holdings-Datensatz verknüpft (1:n-Beziehung)

# Management von Titel- und Bestandsdaten in FOLIO

Im Inventory kommen  
bibliografische Informationen  
zusammen

- Instance                      Titeldaten
- Holdings/Items              Bestandsdaten

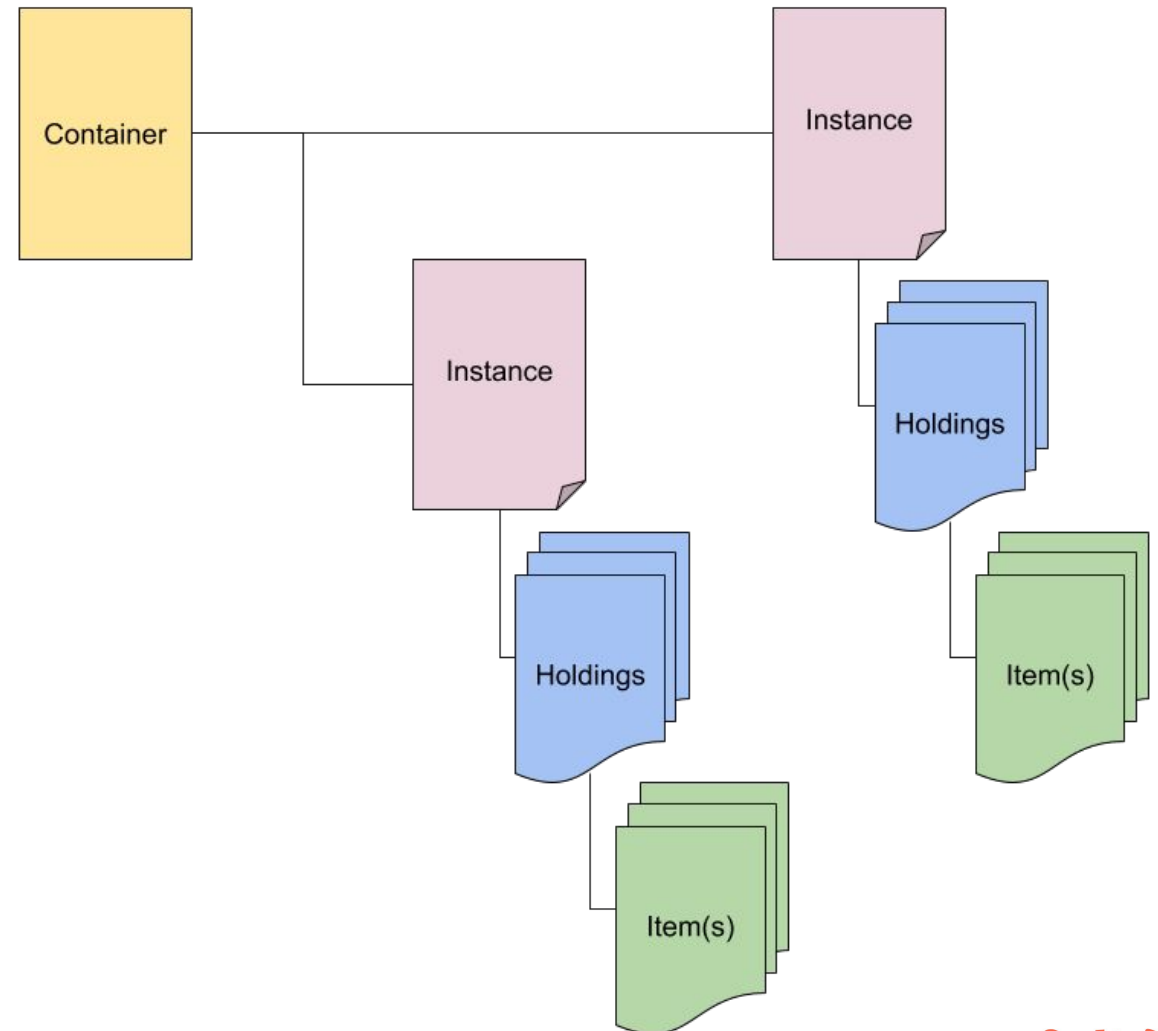




# Management von Titel- und Bestandsdaten cont.

Im Inventory kommen bibliografische Informationen zusammen

- Container: enthalten Instances/Holdings/Items
- Instance: Titeldaten
- Holdings/Items: Bestandsdaten



# Management von Titel- und Bestandsdaten cont.

- Instance-Datensätze != vollständige Katalogisate
- [Inventory-Datenformat](#) enthält auf Instance-Level 61 Elemente
- Zum Vergleich: [Pica3/+ \(K10Plus\)](#) Level 0: > 250 Elemente
- Inventory stellt anderen FOLIO-Apps grundlegende bibliografische Informationen zur Verfügung
- Komplette Katalogisierung findet extern statt



Sehen Sie Bedarf, in FOLIO zu katalogisieren?

# Instance-Schema und Beispiel (JSON)

```
// 20190409110542
// https://raw.githubusercontent.com/folio-org/mod-inventory-storage/master/ramls/instance.json
```

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "description": "An instance record",
  "type": "object",
  "properties": {
    "id": {
      "type": "string",
      "description": "The system assigned unique ID of the instance record; UUID"
    },
    "hrid": {
      "type": "string",
      "description": "The human readable ID, also called eye readable ID. A system-assigned sequential I"
    },
    "source": {
      "type": "string",
      "description": "The metadata source and its format of the underlying record to the instance record"
    },
    "title": {
      "type": "string",
      "description": "The primary title (or label) associated with the resource"
    },
    "indexTitle": {
      "type": "string",
      "description": "Title normalized for browsing and searching; based on the title with articles rem"
    },
    "alternativeTitles": {
      "type": "array",
      "description": "List of alternative titles for the resource (e.g. original language version title"
      "items": {
        "type": "object",
        "properties": {
          "alternativeTitleTypeId": {
            "type": "string",
            "description": "ID for an alternative title qualifier"
          },
          "alternativeTitle": {
            "type": "string",
            "description": "An alternative title for the resource"
          }
        }
      }
    },
    "uniqueItems": true
  },
  "editions": {
    "type": "array",
    "description": "The edition statement, imprint and other publication source information",
    "items": {
      "type": "string"
    },
    "uniqueItems": true
  },
  "series": {
    "type": "array",
    "description": "List of series titles associated with the resource (e.g. Harry Potter)",
    "items": {
      "type": "string"
    },
    "uniqueItems": true
  },
  "identifiers": {
    "type": "array",
    "description": "An extensible set of name-value pairs of identifiers associated with the resource"
    "minItems": 0,
    "items": {
```

```
    "instances": [ {
      "id": "8f3c177d-2511-4469-a58e-8302b85da4fa",
      "hrid": "29301",
      "source": "marc",
      "title": "Canadian journal of statistics.La Revue Canadienne de statistique.",
      "indexTitle": "Canadian journal of statistics.La Revue Canadienne de statistique.",
      "alternativeTitles": [ {
        "alternativeTitleTypeId": "878b6ab7-d20b-4510-bb87-a73f1826dc0c",
        "alternativeTitle": "Canadian journal of statistics"
      }, {
        "alternativeTitleTypeId": "878b6ab7-d20b-4510-bb87-a73f1826dc0c",
        "alternativeTitle": "Revue Canadienne de statistique"
      }, {
        "alternativeTitleTypeId": "878b6ab7-d20b-4510-bb87-a73f1826dc0c",
        "alternativeTitle": "Journal of statistics"
      }, {
        "alternativeTitleTypeId": "878b6ab7-d20b-4510-bb87-a73f1826dc0c",
        "alternativeTitle": "Revue de statistique"
      } ],
      "editions": [ ],
      "series": [ ],
      "identifiers": [ {
        "value": "0319-5724",
        "identifierTypeId": "05c0805a-cf7e-40b1-9095-3405d23cdafa"
      }, {
        "value": "(ICU)BID369062",
        "identifierTypeId": "e07d75a3-9668-40c4-bfdc-3fc083e88d41"
      }, {
        "value": "(OCoLC)1880906",
        "identifierTypeId": "e07d75a3-9668-40c4-bfdc-3fc083e88d41"
      } ],
      "contributors": [ {
        "name": "Statistical Society of Canada."
      }, {
        "name": "Statistical Science Association of Canada."
      } ],
      "subjects": [ "Mathematical statistics--Periodicals.", "Statistics--Periodicals.", "Mathematical statistics.", "Periodicals." ],
      "classifications": [ ],
      "publication": [ {
        "publisher": "Statistical Society of Canada",
        "place": "Nepean, Ont. :",
        "dateOfPublication": "[1973]-",
        "role": "Publication"
      } ],
      "publicationFrequency": [ "Quarterly,<1987- >" ],
      "publicationRange": [ "Vol. 1 (1973)-" ],
      "electronicAccess": [ {
        "uri": "http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)1708-945X",
        "relationshipId": "Version of resource"
      }, {
        "uri": "http://www.jstor.org/journals/03195724.html",
        "publicNote": "Online access via JSTOR:",
        "relationshipId": "Version of resource"
      } ],
    } ],
  }
```



FOLIO	Pica3	Pica+	Beispiel-Inhalt	Kommentar
holdings.dateCreated	E001 OHNE	208@/001 \$a	02-07-18	Feldinhalt: dd-mm-yy : Selektionsschlüssel
holdings.permanent location	7100 \$f	209A/001 \$f	K:	
holdings.illPolicy	7100 \$d	209A \$d	u	Ausleihindikator steuert auch die FL
holdings.updatedByUser	---		---	Mit K10Plus obsolet
item.dateCreated	E001	208@/001 \$a	02-07-18	
item.dateUpdated	7903 OHNE	201B/001 \$0	31-07-18	
item.updatedByUser	---		---	Mit K10Plus obsolet
item.statisticalCode	E0XX OHNE	208@001 \$b	z1k	Selektionsschlüssel
item.statisticalCodeType	---		Selektionsschlüssel	statisticalCode ist ein Key-Value-Pair
item.itemIdentifier	7800 OHNE	203@/001 \$0	1781211647	EPN
item.itemLevelCallNumber	7100 \$a	209A/001 \$a	A 276184	
item.permanentLoanType	7100 \$d	209A \$d	u	
item.permanentHoldingsLocation	7100 \$f	209A/001 \$f	K:	Vererbt vom Holdings-Datensatz
item.permanentItemLocation	7100 \$f	209A/001 \$f	K:	
<i>item.loanDate</i>	---		01-05-19	Ursprung: Ausleih-Apps



# Datenmodell (stark vereinfacht)

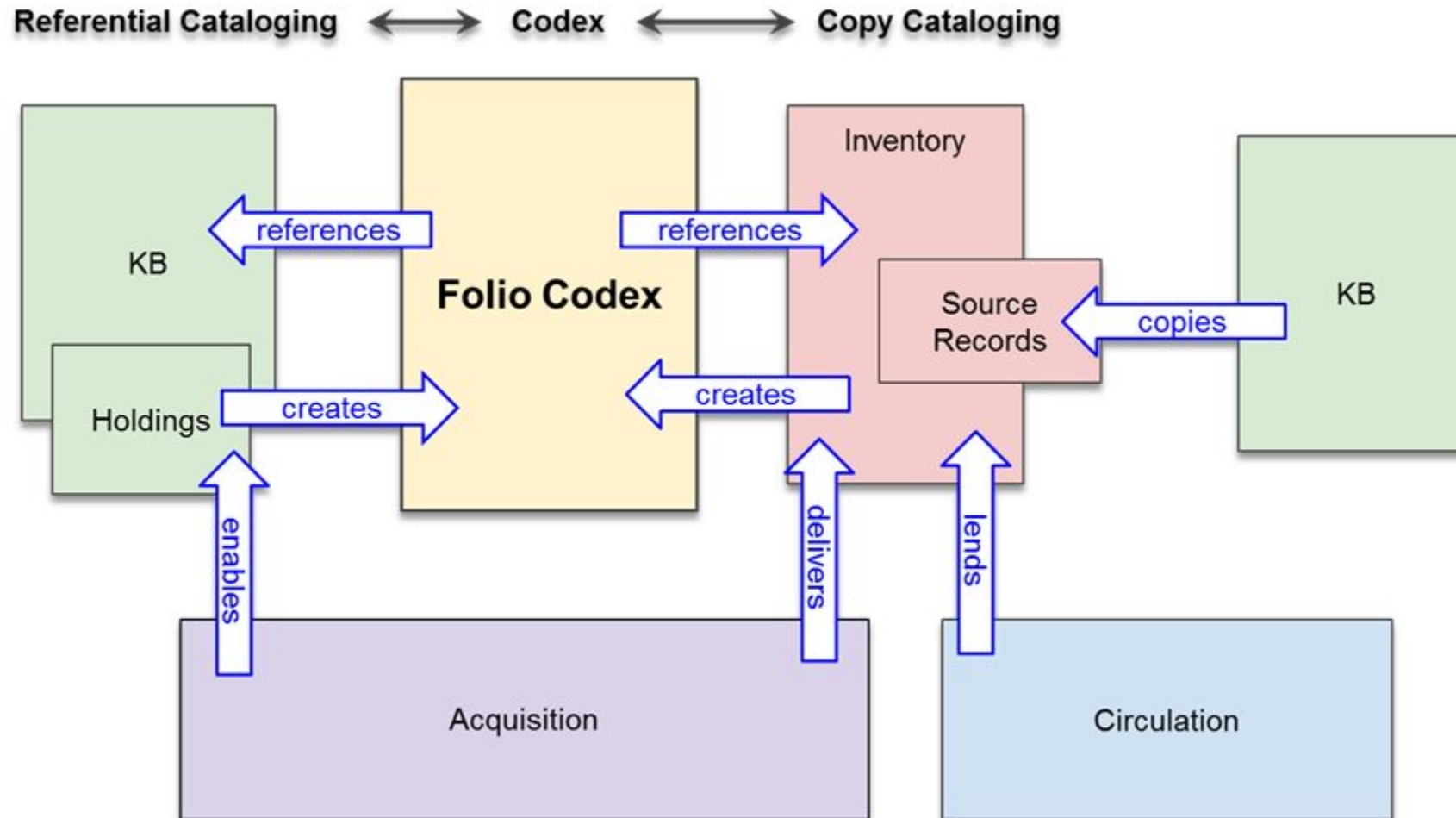


Abbildung aus: Whittenberger, Lynn: Codex Data Objects & Metadata fields : Presentation from FOLIO Forum held on August 16, 2017. 16. August 2017.

# Datenmodell (aktuell)

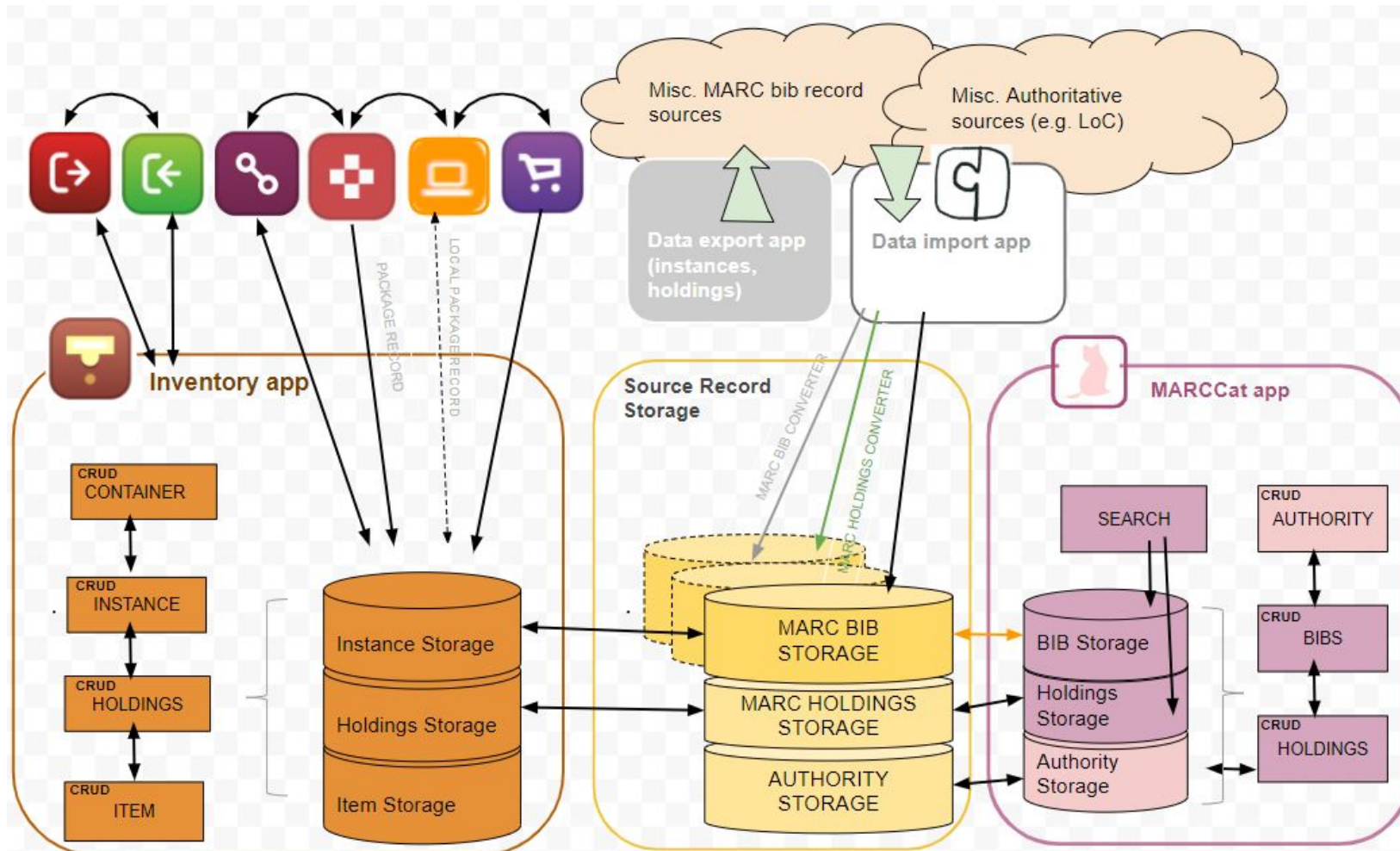


Abbildung aus: <https://docs.google.com/drawings/d/1Vqx5wyPIFIBfckuazSybOgkQgoDDby5y4JDKzJVb24Y/edit>.



# ETL-Prozesse



# ETL-Prozesse



[Auswahl]

Extract

Transform

Load

Auswahl der  
Quelldaten

Datenexport aus dem  
Bibliothekssystem

Konvertierung in  
das Zielformat  
→ Mapping  
★ homogenisieren  
★ bereinigen  
★ anreichern

Datenimport  
→ API  
→ SQL  
→ ...

bibliotheksspezifisch  
(intellektuell)

bibliotheksspezifisch

bibliotheksspezifisch

(zentrale)  
FOLIO-Verfahren

# Migration bibliografischer Metadaten unter technischen Gesichtspunkten



# Datenimport vs. Datenmigration

- **Datenimport**

- Importieren (neu laden oder aktualisieren) oder Löschen eines einzelnen oder weniger Datensätze
- für den täglichen Betrieb, z.B. Austausch Verbund- mit Lokalsystem, Abgleich Nutzerdaten mit Campus-Nutzerverwaltung o.ä.
- Performance ist untergeordnetes Thema

- **Datenmigration**

- Einmaliger, initialer Import von Daten aus dem Vorgängersystem
- große Datenmengen, Performance der Apps ist kritisch
- Wiederholbarkeit und Löschfunktion auf bereits erfolgte Imports notwendig, da während Migrationsphase benötigt
  - Datenkontrolle

# Anforderungen an FOLIO Migration Data Loader

- ein Kommandozeilen-Tool, das JSON-Dateien an alle migrations-relevanten Module über die Modul-API POSTen kann
  - für Titel- und Bestandsdaten kann auch MARC-XML ge-POSTet werden
- die erforderlichen JSON-Schemas werden veröffentlicht und aktuell gehalten
  - es gibt ein Tool zur Validierung der transformierten Exportdaten
- Performanz erlaubt tägliches Laden, Testen & Korrigieren
  - Richtwert max. 8 Std. für Laden eines kompletten Datenabzuges
- kann Dateien größer als Arbeitsspeicher der Maschine laden (z.B. ~ 10 Mio. Instance JSON Records)
- Loader schreibt Log Files mit detaillierten Infos über fehlgeschlagene Record-IDs, und allg. Ladestatistiken
- Gleiche Lade-Optionen und Output-Formate für alle Module
- UUIDs, User-IDs und Datumsstempel können von außen vorgegeben werden

# Anforderungen an FOLIO Data Import App

- Unterstützt CRUD in SRS, Inventory, MarcCat und ACQ
- UI - Bedienoberfläche
  - für Matching, Aktionen, Data-Mapping für den Import
  - Sequenzierung (der Aktionen) in Jobs
- Matching: Abgleich zwischen Input- und schon vorhandenen Datenelementen
- Preview-Modus für interaktive Jobs, Möglichkeit zu Editieren vor finalem Commit
- Log-Dateien wie bei Migrations-Lader; Re-Run möglich; In-App Load Reports
- unterstützt MARC21, MARC-XML, EDIFACT, “flache Datenstrukturen”
- Versionskontrolle der Ausgangsdateien im “Source Record Storage”, Transformation ins FOLIO Internformat für den Modul-Speicher
- Automatische Einplanung (Scheduling) von Jobs, “Pickup” von Importdateien, Job-Profile (v2)



# APIs und Apps für Datenmigration (1)

## mod-data-loader

/load/marc-data/      lädt MARC binär nach Inventory Storage

/load/marc-data/test      Konversion MARC => Instance JSON

Mapping MARC => FOLIO Instance in "Rules"-Datei

Z.Zt. nicht Bestandteil des Releases

Zusammenführung mit Data Import App geplant

# APIs und Apps für Datenmigration (2)

**FOLIO Console Application** <https://github.com/jemiller0/Folio>

Open Source Migrationstool entwickelt an der UChicago

load, save, delete mit Query-Funktionen

- “where”-Bedingung frei programmierbar in SQL oder CQL

Systeminteraktion über die APIs oder direkt in SQL (Postgres)

Äußerst performant und leicht bedienbar

- lädt 100.000 Benutzer in 13 Sek.

Gebaut mit Microsoft Visual Studio - läuft mit .NET / dotnet unter Windows, MAC und Linux

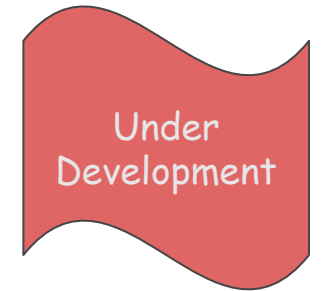
# APIs und Apps für Datenimport : Titel- und Bestände (1)

mod-data-import

mod-source-record-storage, mod-source-record-manager

mod-marccat (FOLIO Katalogisierungsmodul; Titelsätze in MARC21, Holdings in MFHD)

- lädt Datenexporte (z.B. MARC) in den “Source-Record-Storage”
- von dort automatische Überführung ins FOLIO Format +  
in den Modul-Speicher (geplant Q2-2019)
- für Bib-Daten → Inventory, Erwerbungsdaten → ACQ App
- nicht zum Laden von Benutzerdaten, ERM, Ausleihdaten usw.



# Data Import App Details

## Source Record Storage (SRS) vs. Inventory:

- falls MARC vorliegt, ist dieses maßgeblich. Datensatz wird dann in MARCcat editiert, Editieren des Inventory Instance Record ist geblockt.
- Ebenso Synchronisation von SRS/Marccat MARC Holdings und Inventory Holdings; geplant für Q3-2019

## Data Import App Geschwindigkeit

- $45.000 \text{ Datensätze} / 60 \text{ sek} = 750 \text{ Datensätze/sek}$  beim Laden von Marc Bibs in den SRS

# APIs für den Datenimport - Titel und Bestandsdaten (2)

## mod-inventory-storage

Speichert einzelne Datensätze in der Speicherschicht des Inventory-Moduls (Postgres-basiert)

Endpoints (jeweils POST, PUT, GET [erlaubt CQL-Suche], DELETE) für:

- instance-storage - lädt auch MARC JSON
- holdings-storage
- items-storage
- und viele weitere (z.B. instanceRelationship, service-points ...)

# APIs für den Datenimport - weitere Daten

mod-users-import      Benutzerdatenimport

mod-circulation-storage      Speichermodul für Ausleihen

- Ausleihdatensätze, Ausleihregeln

mod-orders-storage      Bestellungen

mod-agreement, mod-licenses      ERM-Module; License Agreements

und viele weitere (z.B. mod-invoices [Rechnungen], mod-finances-storage [z.B. Budgetstrukturen], mod-feesfines [Gebühren])



# (formale) Datentypen im lokalen Bibliothekssystem





# Stammdaten / Referenzdaten

## MODUL- ÜBERGREIFEND

### Kennungen

→ Befugnisse

### *Bibliograf. Metadaten*

→ Ausleihindikator

→ Materialart

## AUSLEIHE: STAMMDATEN

### Nutzerdaten:

Persönliche Daten +

Adressen

## AUSLEIHE: REFERENZDATEN

→ Nutzertypen

→ Kostenarten

→ Nutzerstatus

→ Ausleihtypen

→ Bandstatus

## ERWERBUNG

### Lieferanten

→ Bestelltypen

→ Budgets

...

# Referenzdaten: *Patron groups*

## ^ User information

Last name \*

First name

Patron group \*

Select patron group ▼

Status \*

Active ▼

Select patron group

Alumni (Alumni)

^ Externe / Stadtleser

FL-Bibliothek

FOLIO-Mitarbeiter (keine Ausleihe)

Lehrbeauftragte

Master-Studierende (Graduate Student)

Professoren (Faculty Member)

Studierende (Undergraduate Student)

TVP (Technisches und Verwaltungspersonal)

## ^ User information

Last name \*

First name

Patron group \*

Select patron group ▼

Status \*

Active ▼

Select patron group

10 Studierende (Undergraduate Student)

^ 11 Master-Studierende (Graduate Student)

12 Alumni (Alumni)

20 Professoren (Faculty Member)

25 Lehrbeauftragte

30 TVP (Technisches und Verwaltungspersonal)

40 Externe / Stadtleser

50 FL-Bibliothek

99 FOLIO-Mitarbeiter (keine Ausleihe)

# Transaktionsdaten: Ausleihe

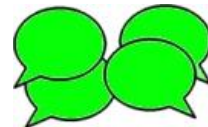
## Daten

- Ausleihen
- Vormerkungen
- (Magazin-)Bestellungen
- Forderungen

wollen wir diese  
Daten migrieren?

## Überlegungen → TO DO

- Zukünftiger Workflow?
- (Wie) passen die Daten in die FOLIO-Welt?
  - Aktuelle Ausleihpolitik “übersetzen”
    - Anderes Konzept?!?
    - Rechtliche Grundlagen beachten (Gebührenordnung, Bibliotheksordnung)



# Sonderfall: Historische Transaktionsdaten

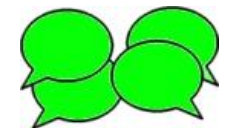
## Ausleihe

- LBS: sehr wenige gespeicherte Informationen (Ausleihgeschichte), wenn überhaupt
- LBS: keine Gebührenhistorie
- Andere Bibliothekssysteme:  
Sind solche Daten vorhanden? Sollen diese Daten migriert werden?

## Erwerbung

- Erledigte/abgeschlossene Bestellungen migrieren?
- Gekündigte/abgeschlossene Abonnements migrieren?
- Erwerbsdaten als Zugangsbuch → ist eine Liste dieser abgeschlossenen Erwerbungen (juristisch/inhaltlich) ausreichend?

Alternativ zum Import in FOLIO:   Laden ins Data Warehouse?



# User-Daten in FOLIO





# User-Daten in FOLIO

Wie gehen wir damit um?

Idee → Users App:  
Datenfelder prüfen



Bibliotheksnutzer

Patrons / Borrower /  
Entleiher

(Mitarbeiter-)  
Kennungen



# Users App

## Bearbeitungsschirm

### TestDemo, Dina

[Collapse all](#)

#### ^ User information

Record last updated: 3/17/2019 5:30 PM

Last name *	First name	Middle name	Barcode
<input type="text" value="TestDemo"/>	<input type="text" value="Dina"/>	<input type="text"/>	<input type="text" value="00180012345"/>
Patron group *	Status *	Expiration date	
<input type="text" value="Studierende (Undergra"/>	<input type="text" value="Active"/>	<input type="text" value="2020-03-30"/>	

#### ^ Extended information

Date enrolled	External system ID	Birth date	Folio number
<input type="text" value="2019-03-17"/>	<input type="text"/>	<input type="text" value="YYYY-MM-DD"/>	<input type="text" value="ce9c8745-d5e0-4fc9-a6ff-c9972bf63602"/>
Username	Folio password		
<input type="text"/>	<a href="#">Send reset password email</a>		

#### ^ Contact Information

Email	Phone	Mobile phone	Preferred contact *
<input type="text" value="klute@gbv.de"/>	<input type="text"/>	<input type="text"/>	<input type="text" value="Email"/>

#### Addresses

**Address1** ☐ Use as primary address

Address Type	Address line 1	Address line 2	City
<input type="text" value="Home"/>	<input type="text" value="Hamburger Straße 22"/>	<input type="text"/>	<input type="text" value="Hamburg"/>
State/Prov./Region	Zip/Postal Code	Country	
<input type="text"/>	<input type="text" value="21073"/>	<input type="text" value="Select Country"/>	

# Daten in der Users App

## Kennungen

- im LBS: Kennung, Name bzw. Bezeichnung, Befugnisse je Modul
- Namensvarianten  
(Uschi Klute / Klute, Uschi / Frau Klute / Thekenplatz 1 / ...)
- LBS-Befugnisse gelten für LBS-Workflows  
→ nicht übertragbar

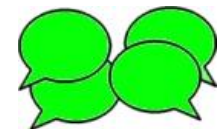
## Nutzerdaten

- vermutlich migrieren  
(Mapping/Konvertierung)
- ggf. vorher alte Datensätze löschen (inaktive Nutzer)
- Datenfelder bereinigen  
(Inhalte ergänzen / löschen / korrigieren / vereinheitlichen)
- Export aus IDM oder ILS?

Schlussfolgerungen?

# Daten in der User App - Fragen

- FOLIO-Login per Username → Feld muss gefüllt sein  
Bisherige Kennungen als Username übernehmen? Oder neue Struktur?
- FOLIO-Prinzip: persönliche Kennungen  
Wollen wir (auch) zukünftig Funktionskennungen verwenden?
- Feld `type` im JSON-Datenschema (*The class of user*)  
zur Unterscheidung von Mitarbeiterkennungen und Nutzerdaten gedacht? Welche funktionalen Abhängigkeiten wären sinnvoll?
- Username auch für Bibliotheksnutzer (=patrons)?  
kein Systemlogin nötig, aber Zugriff auf Konto im Discoverysystem!?  
Ausweisnummer als Username?
- Zwei Accounts für Mitarbeiter, die auch ausleihen wollen?



# Migration der Nutzerdaten





# Users: Datenschema (JSON)

```
"barcode": {  
  "description": "The library  
    barcode for this user",  
  "type": "string"  
},  
"patronGroup": {  
  "description": "A UUID  
    corresponding to the group  
    the user belongs to",  
  "type": "string"  
},  
"lastName": {  
  "description": "The user's  
    surname",  
  "type": "string"  
},
```

```
"firstName": {  
  "description": "The user's  
    given name",  
  "type": "string"  
},  
"email": {  
  "description": "The user's  
    email address",  
  "type": "string"  
},  
"expirationDate": {  
  "description": "The date when the  
    user becomes inactive",  
  "type": "string",  
  "format": "date-time"  
},
```



# Datenanalyse

- Warum ist das wichtig?
  - genaue Kenntnis → Entscheidung, welche Felder migriert werden
  - → Mapping / Konvertierungsvorgaben
  - → Datenbereinigung
- Was muss ich tun?
  - Datenfelder inhaltlich und formal/syntaktisch prüfen
- Wie mache ich das?
  - abhängig von den Zugriffsmöglichkeiten im Bibliothekssystem
- Wer macht das?
  - fachliche Experten des aktuellen ILS / ggf. mit IT-Unterstützung

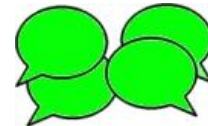


# Mapping: Nutzerdaten

FOLIO	LBS	Beispiel-Inhalt	Kommentar
barcode	borrower.borrower_bar	00180012345	
username	???		für den Zugriff auf das Konto per Discovery-System?
patronGroup	borrower.borrower_type	30	auch Beschreibung aus lbs_code_desc berücksichtigen
lastName	borrower.name	Klute	
middleName	---	---	
firstName	borrower.first_name _initials_prefix	Uschi	hier auch Präfixe wie "von"
email	address.email_address	klute@gbv.de	1. Mehrfachvorkommen: - mehrere E-Mails in 1 Feld / mehrere Adressdatensätze 2. Nutzer ohne E-Mail-Adresse?
expirationDate	membership.expiry_date	2012-12-14 00:00:00.0	Was tun bei mehreren Abteilungsgruppen?
---	borrower.gender	1	Weglassen? Als <i>Tag</i> speichern?
---	borrower.person_titles	Dr.	Wo speichern?

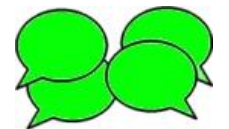
# Mapping: Herausforderungen

- Feld im aktuellen ILS = Feld in FOLIO ✓
- Felder im aktuellen ILS  $\approx$  FOLIO-Felder
- 1:n | n:1 | Freitext vs. normiertes Vokabular | ...
- Feld im aktuellen ILS  $\rightarrow$  kein Feld in FOLIO
- Obligatorisches Feld in FOLIO  $\rightarrow$  kein Feld im Altsystem



# Datenbereinigung 1

- Korrekturen müssen sich am Zielformat orientieren
- Was bereinigen?
  - Konzentration auf Stammdaten, hier vor allem:
    - Erfassungs-/Tippfehler korrigieren
    - Angaben vereinheitlichen / normieren
    - Fehlende Angaben ergänzen
    - Selten genutzte Datenfelder prüfen:  
Eingabe im falschen Feld? “Früher” übliche Erfassung?
  - z. T. auch Transaktionsdaten, wenn diese migriert werden  
(z. B. Forderungen 0,01 €)



# Datenbereinigung 2

- Wie bereinigen?
  - abhängig vom aktuellen ILS und den Zugriffsmöglichkeiten
    - LBS: Für Korrekturen kleinerer Datenmengen online im LBS-Client;  
Falls technisch/inhaltlich möglich: Bereinigung per SQL direkt in der LBS-Datenbank
- Wer bereinigt?
  - fachl. Experten | IT | Personal der betr. Fachabteilung
- Wann bereinigen?
  - jetzt...
    - nützlich auch im aktuellen ILS



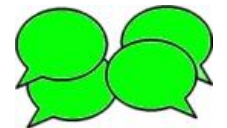
# Datenbereinigung - Beispiele

Nutzerdaten: regelmäßig alte, nicht mehr aktive Konten löschen

- DSGVO
- Vorteil im aktuellen ILS bei der Nutzerdatensuche
- Vorteil bei Datenanalyse: Nur aktuell verwendete Datenfelder müssen berücksichtigt werden (keine Altlasten: „früher haben wir immer ... erfasst“)

Kennungen: regelmäßig die Angaben auf Aktualität kontrollieren

- Ausgeschiedene/neue Mitarbeiter
- Namensänderungen
- Kontrolle der Befugnisse (LBS: Einrichten von Befugnisgruppen)







Danke für Ihre Aufmerksamkeit!

Felix Hemme

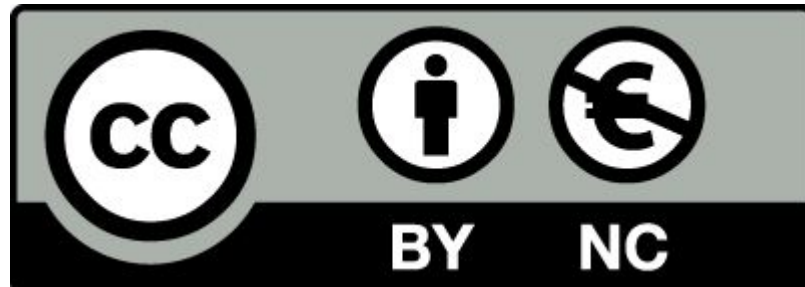
[f.hemme@zbw.eu](mailto:f.hemme@zbw.eu)

Uschi Klute

[klute@gbv.de](mailto:klute@gbv.de)

Ingolf Kuss

[kuss@hbz-nrw.de](mailto:kuss@hbz-nrw.de)



Der Text dieser Präsentation wird unter der Lizenz Creative Commons Namensnennung-Nicht kommerziell 4.0 International (CC BY-NC 4.0) veröffentlicht: <https://creativecommons.org/licenses/by-nc/4.0/>

Davon ausgenommen sind die verwendeten, nicht von den Autoren erstellten Grafiken, Screenshots und Bilder, deren jeweilige Rechte und Lizenzbedingungen fortgelten.

Maßgeblich für diese Präsentation ist das gesprochene Wort.



# LBS: Nutzerdaten (Tabelle borrower)

address_id_nr	111760
borrower_bar	22123458
iln	22
reminder_address	1
ill_number	0
name_key	TESTKU
registration_number	22123458
borrower_status	0
borrower_type	40
date_of_birth	03.09.1964 00:00
edit_date	2019-03-23 12:27:17.053
first_name_initials_prefix	Uschi
free_text	zahlt Mahngebühren nur nach langen Diskussionen
gender	2
input_date	26.01.2001 13:23

language_code	DE
message	Bitte in der Bibliothek melden!
name	TestKlute
person_titles	Dr.
proof_of_identity	L2FKPJ34H
query_name	TESTKLUTE,USCHI
source	1
sub_query_name	TES
borrower_ts	00000000ea16e987
free_text_block	(null)
email_address	(null)
institution_code	0
date_time_last_active	17.03.2019 19:07
pica_timestamp	(null)
end_user_key	(null)
no_log_action_idm	0



# LBS4-Client: Nutzerstammdaten

**Ausleihe [Theke 1 - Ausleihzentrum]** Benutzer anzeigen/ändern

Benutzernr.	<input type="text" value="0022123458"/>	Benutzertyp	<input type="text" value="40"/> TVP	<b>Anmerkung</b> zahlt Mahngebühren nur nach langen Diskussionen	
Registriernr.	<input type="text" value="0022123458"/>	Benutzerstatus	<input type="text" value="0"/> Ausleihe erlaubt		
Namenschlüssel	<input type="text" value="TESTKU"/>	Sprachcode	<input type="text" value="DE"/> Deutsch		
Titel	<input type="text" value="Dr."/>	Geburtsdatum	<input type="text" value="03-09-1964"/>		
Vorname	<input type="text" value="Uschi"/>	Nachname	<input type="text" value="TestKlute"/>		
Ausweisinfos	<input type="text" value="L2FKPJ34H"/>	Geschlecht	<input type="radio"/> männlich <input checked="" type="radio"/> weiblich <input type="radio"/> Unbekannt		
Neues Passwort	<input type="text"/>	Ablaufdatum	<input type="text" value="15-03-2022"/>		
Eingabedatum	26-01-2001 13:23	Änderungsdatum	31-03-2019 10:45	Letzte Aktivität	17-03-2019 19:07

Adr   Texte   FL   Ausw   Lokal   €   € (2)   Entl.   Vorm   B-Nr

Ty.	Zimmer/Straße/Postf.	PLZ	Ort	Tel.-Nr.	E-Mail	
	✉ Von-Melle-Park 3	20146	Hamburg	040 / 42838-6377	klute@gbv.de	✉
	Beamtenlaufbahn 7	22846	Norderstedt	040 / 123 4567	uk@web.de	✉



# LBS4-Client: Kommentarfelder

Adr

Texte

FL

Ausw

Lokal

€

€ (2)

Entl.

Vorm

B-Nr

Interne Bemerkg.

zahlt Mahngebühren nur nach langen Diskussionen

Mitteilung

Bitte in der Bibliothek melden!



# LBS4-Client: Adressdaten

Adresse ändern -- Webseitendialog

Name	TestKlute Uschi	Adresstyp	
<u>Z</u> immer		z. <u>H</u> d.	
<u>S</u> traße/Postf.	Von-Melle-Park 3	<u>P</u> LZ	20146
<u>O</u> rt	Hamburg	<u>c</u> /o	Verbundzentrale des GBV
<u>B</u> undesland		<u>L</u> and	
<u>T</u> elefon	040 / 42838-6377	<u>T</u> elefon2	
<u>F</u> ax		<u>E</u> -Mail-Adresse	klute@gbv.de
<u>A</u> dresszusatz		<u>M</u> ahnadresse	<input checked="" type="checkbox"/>





# LBS4-Client: Lokale Felder

Adr	Texte	FL	Lokal	€	Entl.	Vorm	B-Nr	Gesch.
Standort							W	
Abteilung							Physik des Erdsystems	
Benutzertyp							40	x

Weitere von LBS-Bibliotheken eingerichtete lokale Felder (Benutzerdaten):

Geburtsname, Ort FL-Bibliothek, FL-Sigel, URL, Alte Benutzernummer, Benutzer eingegeben von, Research subject, Vorname, Jahrgang, Studiengang, Online-Portal, Hochschule, ...



# FOLIO: Users - allg. Daten

## ^ User information

▼ Record last updated: 4/2/2019 10:33 AM

Last name *	First name	Middle name	Barcode
<input type="text" value="TestKlute"/>	<input type="text" value="Uschi"/>	<input type="text" value="Wonderwoman"/>	<input type="text" value="22123458"/>
Patron group *	Status *	Expiration date	
<input type="text" value="TVP (Technisches und v"/>	<input type="text" value="Active"/>	<input type="text" value="2022-03-14"/>	

## ^ Extended information

Date enrolled	External system ID	Birth date	Folio number
<input type="text" value="2001-01-25"/>	<input type="text" value="22123458"/>	<input type="text" value="1964-09-02"/>	<input type="text" value="ba999ae4-a119-4bc7-8b6e-c7bd745e7baa"/>
Username	Folio password		
<input type="text" value="22123458"/>	<a href="#">Send reset password email</a>		





# FOLIO: Users - Adressinformationen

## ^ Contact Information

Email	Phone	Mobile phone	Preferred contact *
<input type="text" value="klute@gbv.de"/>	<input type="text" value="040 / 42838-6377"/>	<input type="text"/>	<input type="text" value="Email"/>

## Addresses

<b>Address1</b> <input checked="" type="radio"/> Use as primary address 			
Address Type	Address line 1	Address line 2	City
<input type="text" value="Work"/>	<input type="text" value="Von-Melle-Park 3"/>	<input type="text"/>	<input type="text" value="Hamburg"/>
State/Prov./Region	Zip/Postal Code	Country	
<input type="text"/>	<input type="text" value="20146"/>	<input type="text" value="Select Country"/>	
<b>Address2</b> <input type="radio"/> Use as primary address 			
Address Type	Address line 1	Address line 2	City
<input type="text" value="Home"/>	<input type="text" value="Beamtenlaufbahn 7"/>	<input type="text"/>	<input type="text" value="Norderstedt"/>
State/Prov./Region	Zip/Postal Code	Country	
<input type="text"/>	<input type="text" value="22846"/>	<input type="text" value="Select Country"/>	

# FOLIO: Users App auf deutsch



Nutzer

Nutzer

Bestand

Ausleihe

Rückgabe

Vormerkungen

Codexsuche

ERM

Data Import

eHoldings

stripes-core.mainnav.showAllApplicationsButtonLabel

Search & Filter

test

Suche

Reset all

Status

Include inactive users

Benutzergruppe

Alumni

Externe / Stadtleser

FL-Bibliothek

FOLIO-Mitarbeiter

Lehrbeauftragte

Master-Studierende

Professoren

Studierende

TVP

Visiting Scholar

Nutzer

5 Records found

+ New

Active

Name

Barcode

Active

TestDemo, Dina

0018001234

Active

TestKlute, TestUschi test

Active

TestKlute, Uschi Wonderwoman

22123458

Active

Testnutzerin, Luna

8989346711

Active

TestStudentin, Susi

0018009876

TestKlute, Uschi Wonderwoman

Expand all

Benutzerinformationen

Record last updated: 2.4.2019 10:33

Nachname

Vorname

weitere Vorname(n)

Barcode

TestKlute

Uschi

Wonderwoman

22123458

Benutzergruppe

Status

Ablaufdatum

TVP

active

14.3.2022

Patron Blocks

Erweiterte Informationen

Erfassungsdatum

Geburtsdatum

25.1.2001

3.9.1964

Folio-ID

Externe System-ID

ba999ae4-a119-4bc7-8b6e-c7bd745e7baa

22123458

Benutzername

22123458

Kontaktinformationen

E-Mail

Telefon

Mobiltelefon

Bevorzugte Kontaktmöglichkeit

klute@gbv.de

040 / 42838-6377

E-Mail

Adresse(n)

Primary

Address Type

Address line 1

Address line 2

City

Work

Von-Melle-Park 3

Hamburg

State/Prov./Region

Zip/Postal Code

Country

20146

Show more addresses(1)



# FOLIO: User mit Sperre

Users

Users

Inventory

Check out

Check in

Requests

Codex Search

ERM

Data Import

eHoldings

Apps

Q

Users

5 records found

New

Active	Name	Barcode	Patron group	Username
Active	TestDemo, Dina	00180012345	Studierende	
Active	TestKlute, TestUschi test		FOLIO-Mitarbeiter	test
Active	TestKlute, Uschi Wonderwoman	22123458	TVP	22123458
Active	Testnutzerin, Luna	8989346711	Studierende	
Active	TestStudentin, Susi	00180098765	Studierende	

TestDemo, Dina

Patron has blocks in place

Expand all

User information

Record last updated: 3/17/2019 12:30 PM

Last name	First name	Middle name	Barcode
TestDemo	Dina		00180012345
Patron group	Status	Expiration date	
Studierende	active	3/30/2020	

Patron Blocks

Add Block

Type	Display description	Blocked actions
Manual	keine Bestellungen	Requests

Extended information

Contact Information

Proxy/Sponsor

Fees/Fines

Loans

Requests

User permissions

Service points