

Import Data using the OpenEHR-Flat-Loader

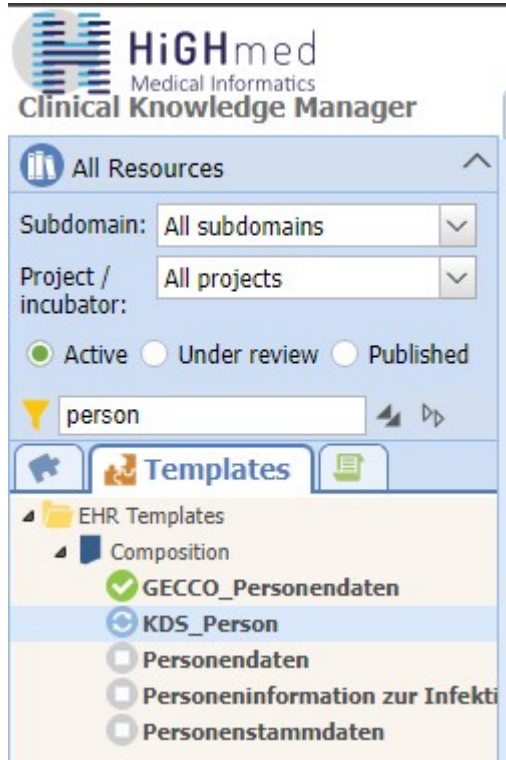
Manual and Tips by Jendrik Richter

- Question to jendrik.richter@med.uni-goettingen.de

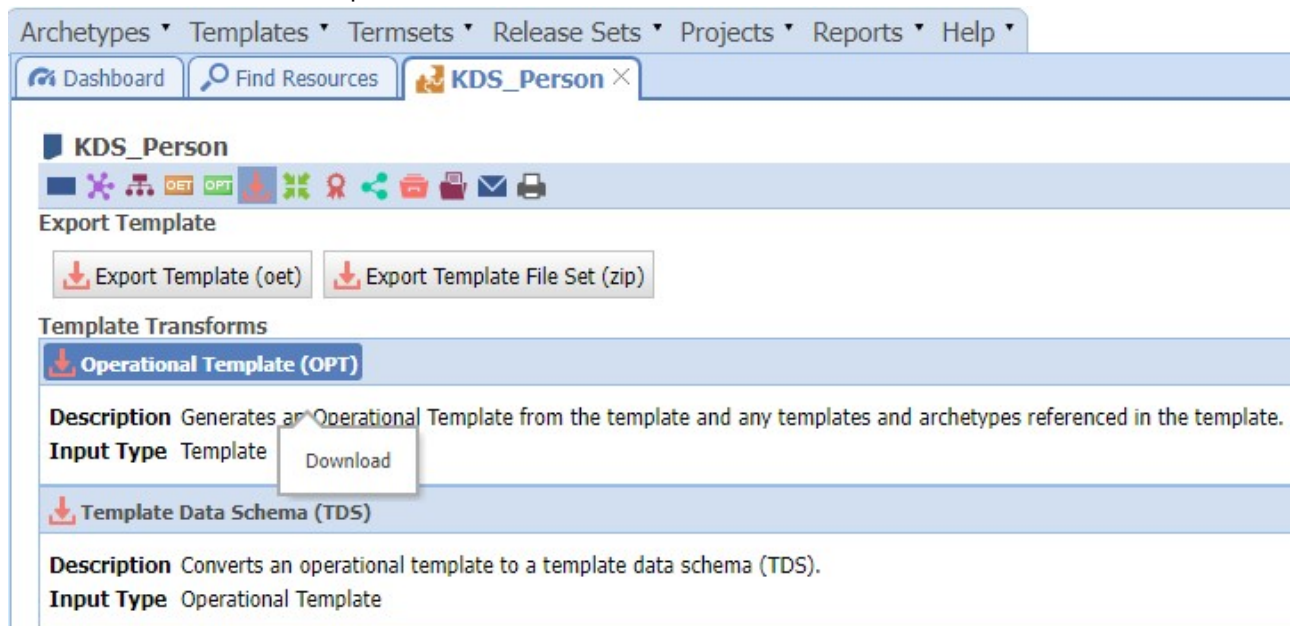
Prerequisite

OPT:

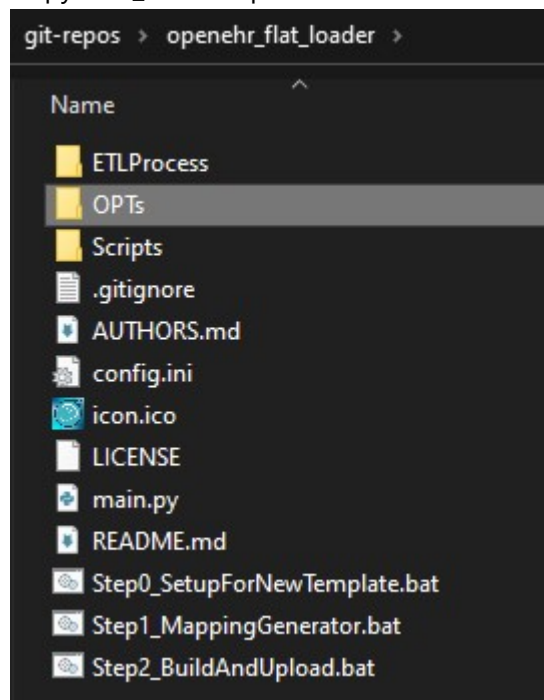
- Search the KDS_Person Template in the [CKM](#)



- Download KDS_Person-Template from [CKM](#)



- Copy KDS_Person.opt to OPTs-Folder of the FLAT-Loader



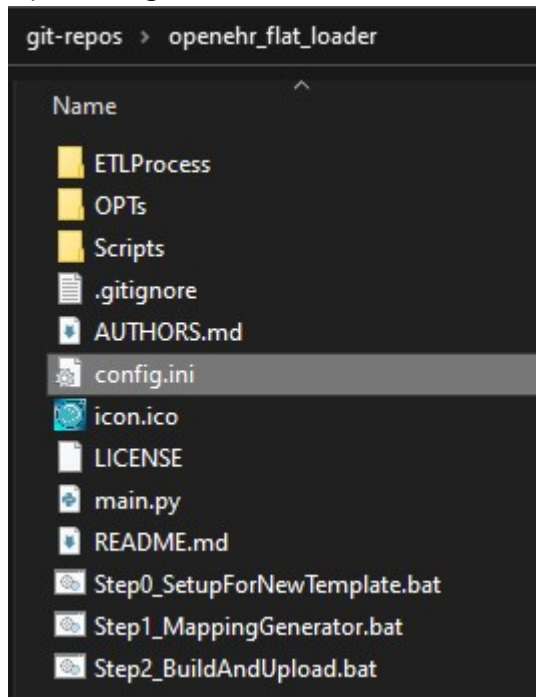
CSV:

- Provide data as CSV-File and copy CSV-File to ETLProcess/Input

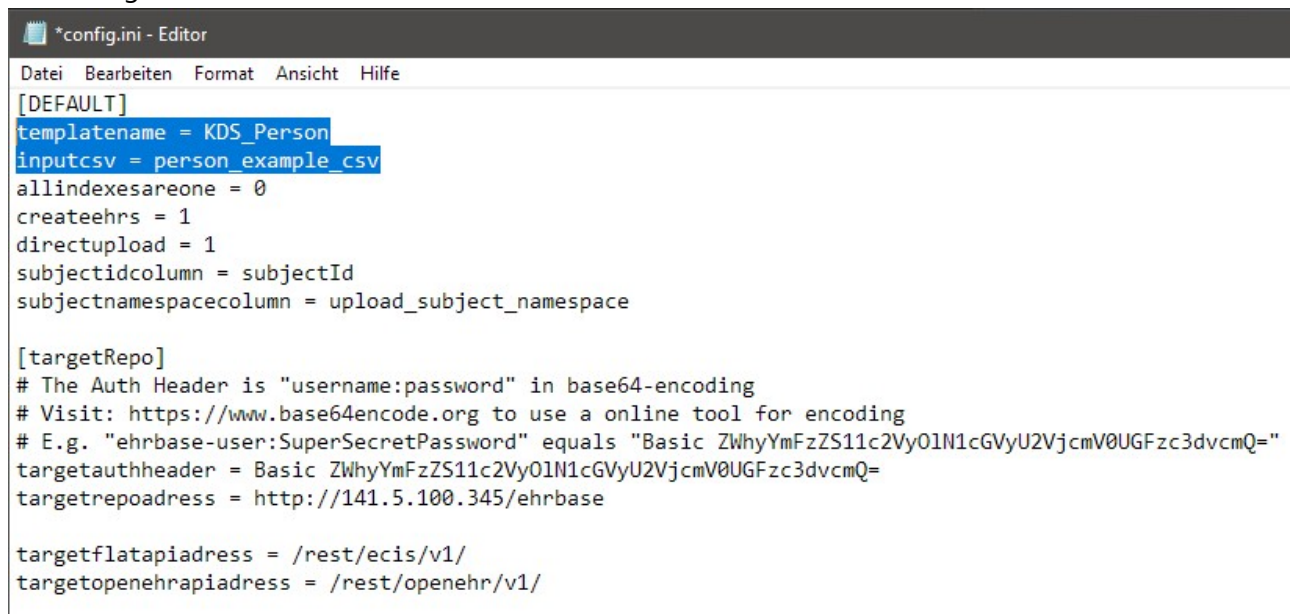
git-repos > openehr_flat_loader > ETLProcess > Input			
Name	Änderungsdatum	Typ	Größe
.gitkeep	28.06.2022 15:22	GITKEEP-Datei	0 KB
jpeg_data.csv	28.06.2022 15:22	Microsoft Office E...	238 KB
NATARS_Tzusatz.csv	28.06.2022 15:22	Microsoft Office E...	3 KB
person_example_csv.csv	18.07.2022 09:51	Microsoft Office E...	2 KB
test1_ansi.csv	28.06.2022 15:22	Microsoft Office E...	3 KB
test1_utf8.csv	28.06.2022 15:22	Microsoft Office E...	3 KB
ucc_score_person_utf8_updated.csv	28.06.2022 15:22	Microsoft Office E...	1 KB

Config:

- Open config-File of the Flat-Loader



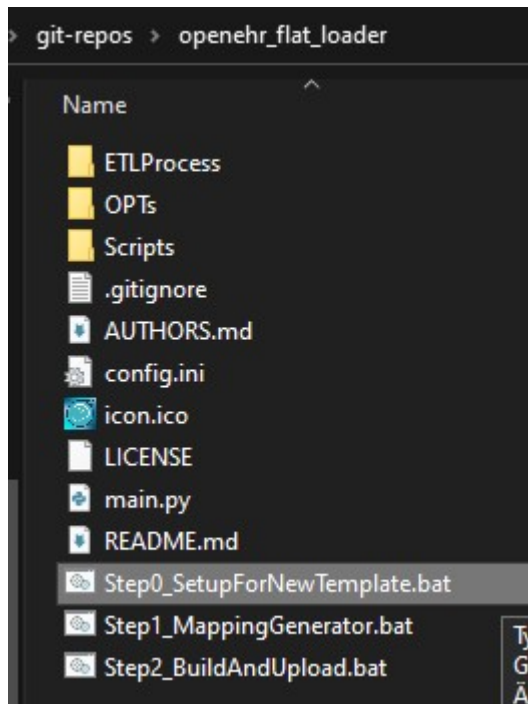
- Edit config-entries to match use case



Person Example

Setup OPT at Server:

- Run "Step0_SetupForNewTemplate.bat" to upload the OPT to your server (defined by the IP in the config)



- The Output may look like this:

```
C:\WINDOWS\system32\cmd.exe
Used Argument: -generateExamples

HandleOPT is running:
OP Exists? 200
    OPT already exists at this server
    OPT exists at server and WebTemplate has been downloaded

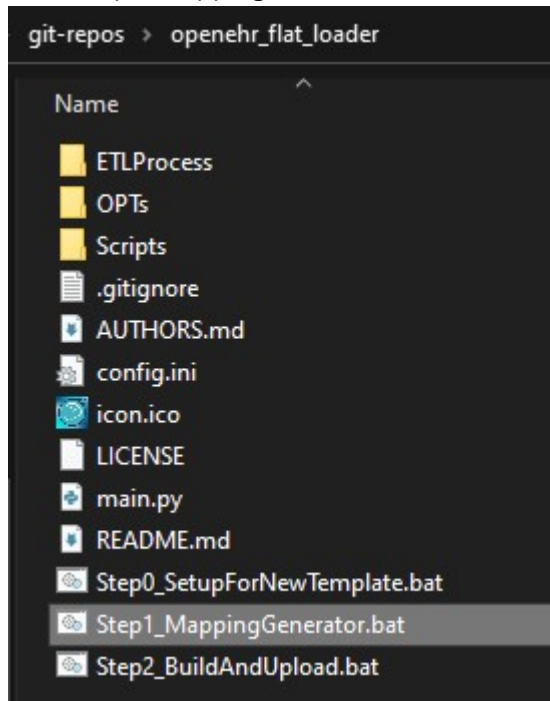
OPT is uploaded to the Repository and an Example-Composition is stored in the ManualTasks-Folder.
```

Make sure necessary data columns are included in CSV:

- Make sure the columns "ehrlId" exists
- Make sure a subjectId-Column exists and is defined in config.
- Make sure a namespace-Column exists and is defined in config.

Generate the Excel-Mapping-Table:

- Run "Step1_MappingGenerator.bat"



- The Output will give you some basic information about the FLAT-Paths that were extracted

```
C:\WINDOWS\system32\cmd.exe
Used Argument: -generateMapping
Welcome to the openEHR_FLAT Loader-Commandline-Tool!
Given an existing template, this tool allows you to transform tabular data into the interoperable openEHR format.
Variables for template, data/csv-file and repository can be specified in config.ini.

HandleOPT is running:
OP Exists? 200
    OPT already exists at this server
    OPT exists at server and WebTemplate has been downloaded
PathExport is running:
    Anzahl extrahierter Pfade: 75
    Extracted FLAT-Paths from the WebTemplate
MappingListGen is running:
    Anzahl der Pflichtpfade (ohne Suffixe): 10
```

- For some paths the template takes several entries (cardinality != ([1..1] or [0..1])).
 - For this entries provide the number of maximum repetitions of the item in your data set.
- In the case of "KDS_Person"-Template all the repeatable items for "Straßenanschrift", "Postfach" and others are queried.
 - Enter either 0 or 1 for this example.
- The Mapping-Generation finishes with generating the empty Mapping-Table

```
Das nachfolgende Element kann in der Composition beliebig oft wiederholt werden:
person/personendaten/person/name/vorname:<<index>>
Wie oft wird das Element maximal pro Composition vorkommen?
2

Das nachfolgende Element kann in der Composition beliebig oft wiederholt werden:
person/personendaten/kontaktperson/organisation/funktion:<<index>>
Wie oft wird das Element maximal pro Composition vorkommen?
1
Generated the (empty) Mapping-Table
```

Perform the Mapping (and enrich the Data-CSV)

- You can find all related files under "openehr_flat_loader\ETLProcess\ManualTasks"

git-repos > openehr_flat_loader > ETLProcess > ManualTasks >

Name	Änderungsdatum
jpeg_bild	28.06.2022 15:16
KDS_Medikationseintrag	28.06.2022 15:16
KDS_Person	18.07.2022 10:44
Person	28.06.2022 15:16
WebTemplate_Datatypes	15.07.2022 09:24
ZLG_Testdaten	28.06.2022 15:16
.gitkeep	09.06.2022 13:43

- The example composition that has been generated by the server gives a good idea how FLAT-Paths/Items of the Template may be mapped to data items/values

git-repos > openehr_flat_loader > ETLProcess > ManualTasks > KDS_Person

Name	Änderungsdatum	Typ	Größe
KDS_Person_MAPPING.xlsx	18.07.2022 10:44	Microsoft Office E...	13 KB
KDS_Person_WebTemplate.json	18.07.2022 10:29	JSON-Datei	161 KB
KDS_PersonCompositionExample.json	18.07.2022 10:29	JSON-Datei	11 KB

KDS_PersonCompositionExample.json

```

1 {
2   "person/category|value": "event",
3   "person/category|terminology": "openehr",
4   "person/category|code": "433",
5   "person/context/patient_id|id": "dev/null",
6   "person/context/start_time": "2022-02-03T04:05:06",
7   "person/context/setting|value": "home",
8   "person/context/setting|code": "225",
9   "person/context/setting|terminology": "openehr",
10  "person/context/_end_time": "2022-02-03T04:05:06",
11  "person/context/_health_care_facility|name": "DOE, John",
12  "person/geschlecht/administratives_geschlecht|code": "42",
13  "person/geschlecht/administratives_geschlecht|terminology": "http://fhir.hl7.org/ValueSet/$expand?url=http://fhir.de/ValueSet/gender-other-de",
14  "person/geschlecht/administratives_geschlecht|value": "No example for terminology 'http://fhir.hl7.org/ValueSet/gender-other-de'",
15  "person/geschlecht/anderes_geschlecht|value": "divers",
16  "person/geschlecht/anderes_geschlecht|terminology": "http://fhir.de/ValueSet/gender-other-de",
17  "person/geschlecht/anderes_geschlecht|code": "D",
18  "person/geschlecht/language|terminology": "ISO_639-1",
19  "person/geschlecht/language|code": "de",
20  "person/geschlecht/encoding|terminology": "IANA_character-sets",
21  "person/geschlecht/encoding|code": "UTF-8",
22  "person/geschlecht/_uid": "9ea96af0-46cd-3119-a071-772db05a9cf4",
23  "person/geschlecht/_work_flow_id|id": "09b7435f-aebb-3963-850f-6520eb10266e",
24  "person/geschlecht/_work_flow_id|id_scheme": "scheme",
25  "person/geschlecht/_work_flow_id|namespace": "unknown",
26  "person/geschlecht/_work_flow_id|type": "ANY",

```

- You can use the KDS_Person_WebTemplate.json to lookup paths and cardinalities as well as value sets.

- The FLAT-Paths result from the tree-structure of the elements identified by their "id"

```
1  {
2    "templateId": "KDS_Person",
3    "version": "2.3",
4    "defaultLanguage": "de",
5    "languages": [
6      "de"
7    ],
8    "tree": {
9      "id": "person",
10     "name": "Person",
11     "localizedName": "Person",
12     "rmType": "COMPOSITION",
13     "nodeId": "openEHR-EHR-COMPOSITION.person.v0",
14     "min": 1,
15     "max": 1,
16     "localizedNames": {
17       "de": "Person"
18     },
19     "localizedDescriptions": {
20       "de": "Dokument zur Übermittlung von Informationen über Personen"
21     },
22     "children": [
23       {
24         "id": "category",
25         "name": "category",
```

- You can lookup the structure, cardinalities and value sets online in the [CKM](#)

PERSON

Collapse All **Show Annotations** **Show Paths**

Other context

ID Patient ID

Geschlecht [1..1]

data

Administratives Geschlecht [1..1] [administrative-gender]

Anderes Geschlecht amtlich [http://fhir.de/ValueSet/gender-other-de]

Personendaten

data

Person [1..1]

Name

Namensart [1..1] [name-use]

Vollständiger Name

Vorname [0..*]

Familienname [1..1]

Familienname-Namenszusatz

Familienname-Nachname

Familienname-Vorsatzwort

Prefix [0..*]

Suffix [0..*]

- With this documents as support you may be ready to perform the mapping task
 - Open the Mapping-Table in Excel

git-repos > openehr_flat_loader > ETLProcess > ManualTasks > KDS_Person

Name	Änderungsdatum	Typ	Größe
KDS_Person_MAPPING.xlsx	18.07.2022 10:44	Microsoft Office E...	13 KB
KDS_Person_WebTemplate.json	18.07.2022 10:29	JSON-Datei	161 KB
KDS_PersonCompositionExample.json	18.07.2022 10:29	JSON-Datei	11 KB

- Select each CSV-Element that shall be mapped to a Path

A	B	C	D
Index(e)	FLAT-Path (Data field in later composition - If mapped)	Map CSV-Column to Path (Dropdown-Selector)	
1			
2	P	person/category value	
3	P	person/category code	
4	P	person/category terminology	
5		person/context/patient_id id	subjectid
6		person/context/patient_id type	
7		person/context/patient_id issuer	
8		person/context/patient_id assigner	
9	P	person/context/start_time	
10	P	person/context/setting value	
11	P	person/context/setting code	
12	P	person/context/setting terminology	
13	P	person/geschlecht/administratives_geschlecht value	
14	P	person/geschlecht/administratives_geschlecht code	
15	P	person/geschlecht/administratives_geschlecht terminology	

Dropdown menu options: index, subjectid, pidi, ehrfd, upload_subject_namespace, ehr_start_time_iso, geb_dat, gender_person_code_value

- Cells with RED-Colour are mandatory for the composition. (Sometimes they are filled by the server if left blank.)
- Cells with ORANGE-BROWN-Colour are mandatory if the element which they belong to is present.
- If you are missing a CSV-Column (e.g. you have a language but you are missing the terminology-name you have to add it in your CSV-File (per subject) or you have to set it in the Mapping via Column E - "Set Metadata directly (optional)".

Note that values given in Column E are set for every subject/composition)

- Metadata set per subject in the CSV:

T	U	V	W	X	Y	Z
person_name	territory_code	territory_terminology	language_code	language_terminology	bericht_composer_name	
http://hl7.org	DE	ISO_3166-1	de	ISO_639-1	openEHR_Learner	
http://hl7.org	DE	ISO_3166-1	de	ISO_639-1	openEHR_Learner	
http://hl7.org	DE	ISO_3166-1	de	ISO_639-1	openEHR_Learner	
http://hl7.org	DE	ISO_3166-1	de	ISO_639-1	openEHR_Learner	
http://hl7.org	DE	ISO_3166-1	de	ISO_639-1	openEHR_Learner	

- Metadata for territory set directly in the mapping for every subject but metadata for language taken from respective columns in the CSV:

	A	B	C	D	E
1		Index(e)	FLAT-Path (Data field in later composition - if mapped)	Map CSV-Column to Path (Dropdown-Selector)	Set Metadata directly (optional)
141	P		person/composer/name	bericht_composer_name	
142	P		person/language/code	language_code	
143	P		person/language/terminology	language_terminology	
144	P		person/territory/code		DE
145	P		person/territory/terminology		ISO_3166-1
146					
147			Legend:		
148			Mandatory Path that needs to be present to get a valid Composition		
149			Conditionally mandatory Path that needs to be present if the 'parent'-Element is exists		
150			Non-mandatory Path that does not need to be present to store the Composition		

- Note, setting the Values directly for a mandatory field should not result in problems, but setting values directly for non-mandatory field may result in compositions where for example a terminology name is defined but no terminology value or code is present.
- The entries that you need for Terminology-Names, Values from Terminologies, etc. can be looked up in the CKM, WebTemplate, Online, OpenEHR-Specification or the Example-Composition.
- If you want to duplicate a path you have to have an eye on the indexes
 - Make sure index 0 always exists in a data set before also mapping the path for index 1, otherwise the server cant store the resulting compositions.

	A	B	C	D	E
1		Index(e)	FLAT-Path (Data field in later composition - if mapped)	Map CSV-Column to Path (Dropdown-Selector)	Set Metadata directly (optional)
31		0	person/personendaten/person/name/vorname:<<index>>	person_name_vorname	
32		1	person/personendaten/person/name/vorname:<<index>>	person_name_vorname2	

- Todesdiagnose is a great example for enriching of the CSV:

1	A	B	C	D	E
		Index(e)	FLAT-Path (Data field in later composition - if mapped)	Map CSV-Column to Path (Dropdown-Selector)	Set Metadata directly (optional)
106			person/personendaten/angaben_zum_tod/verstorben	verstorben	
107			person/personendaten/angaben_zum_tod/angaben_zum_tod/sterbedatum	sterbedatum_iso	
bP			person/personendaten/angaben_zum_tod/angaben_zum_tod/todesdiagnose value		
bP			person/personendaten/angaben_zum_tod/angaben_zum_tod/todesdiagnose code	todesdiagnose_icd10gm	
bP			person/personendaten/angaben_zum_tod/angaben_zum_tod/todesdiagnose terminology		

- The Mapping requires not only the code from the terminology but also its name and the corresponding value

- In the [CKM](#) the Template has "bfarm/icd-10-gm" defined as the name for our terminology

- We add this as the name of the terminology to our CSV.
- For our subject/entry with "I22.1" as a code from this terminology we now lookup the Display-Value for this code (e.g. using Google).
- This results in the following entries for "Todesdiagnose"

S	T	U
todesdiagnose_icd10gm	todesdiagnose_terminology	todesdiagnose_value
I22.1	bfarm/icd-10-gm	Rezidivierender Myokardinfarkt der

- Note that there is only one subject with entry about death. Therefore it is the right way to supply these metadata directly in the CSV and not as Metadata set directly in the mapping, since subject 1+2+4+5 do not have a diagnosis and therefore should not have any metadata mapped about it.

Caution:

After enriching metadata newly added CSV-columns are not available in the mapping.
You may enter the names directly.

! If you generate new mapping the existing file is overwritten - save your old mapping if needed !

Generating resources and uploading them to the server

- Run "Step2_BuildAndUpload.bat"
- At first the resources are generated and stored in the Output-Folder

```
BuildComp is running.
Erstelle 5 Composition-Ressourcen.
5 / 5 Ressourcen erstellt und im Ordner "Output" gespeichert.
buildComp finished.
```

- After that the EHRs are created at the server

```
Create 5 EHRs:
Hindernis beim EHR erstellen mit Status: 409
EHR existierte bereits mit ehrID: 2ff4c284-a8b1-49e6-af0b-cf44a9a0b3ac

Hindernis beim EHR erstellen mit Status: 409
EHR existierte bereits mit ehrID: a80477cd-54d8-4552-a27d-2f3dd3d24ca9

Hindernis beim EHR erstellen mit Status: 409
EHR existierte bereits mit ehrID: 489ee219-012a-404b-b5d3-470a8a9ed6bb

Hindernis beim EHR erstellen mit Status: 409
EHR existierte bereits mit ehrID: b7aada04-2b38-4ec7-9c81-961b0565e4ee

Hindernis beim EHR erstellen mit Status: 409
EHR existierte bereits mit ehrID: b514ab79-78f9-4988-a942-37f733a16590

EHRs for 5 / 5 Subjects have been created successfully.
```

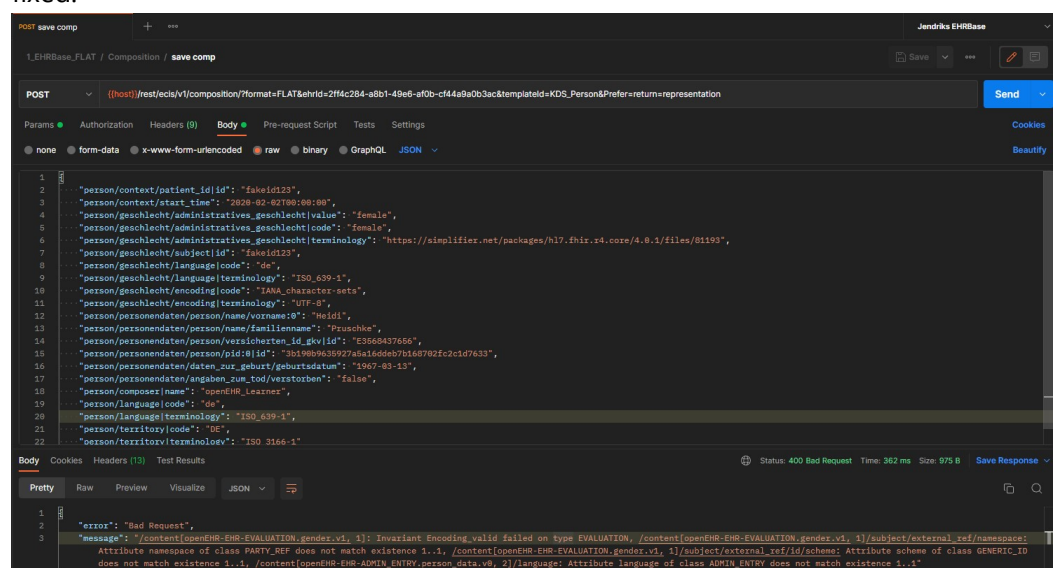
- At last the ressources are uploaded to the server

- This step either succeeds and the ressources are available in the openEHR-Repo.

```
Status beim Upload der Composition: 201
{'meta': {'href': {'url': 'http://141.5.100.115/ehrbase/rest/ecis/v1/composition/?format=FLAT&ehrId=b514ab79-78f9-4988-a942-37f733a16590&templateId=KDS_Person/rest/ecis/v1/composition/088a55d2-1e83-4111-9f7e-229f423051f2::local.ehrbase.org:1'}}}, {'action': 'CREATE', 'compositionUid': '088a55d2-1e83-4111-9f7e-229f423051f2::local.ehrbase.org:1'}
CompositionUid: 088a55d2-1e83-4111-9f7e-229f423051f2::local.ehrbase.org:1

5 / 5 Compositions have been created successfully.
Upload finished. Great Success.
```

- Or you are presented an error message by the server that has to be dealt with.
 - The error messages may range from invalid encodings or missing attributes in a composition to Internal Server Errors.
 - It is often a good process to select some composition that has been created and try to upload it via e.g. Postman until all mistakes in the data-CSV and Mapping are fixed.



- The errors shown above were due to
 - the subjectID being incorrectly mapped to subject|id instead of subject|name
 - one of the language|terminolgy-Paths being mappend to CSV-Column "territory_terminology"

- "namensart" being mapped for the geburtsname-element (person/personendaten/person/geburtsname) instead the name-element (person/personendaten/person/name)
- Please note that after dealing with the errors that were returned with the next request to the server the validation process goes further and more errors may be returned by the openEHR-Repo.

Querying uploaded Compositions

- To access the data in the openEHR-Repository we need to send an AQL-Query via the REST-API
 - Postman is a nice tool to send REST-Request
 - To build an AQL-Request:
 - Select the Item in the Template that you want to query (e.g. Vorname + Nachname)
 - Get the paths and IDs of these items from the CKM
 - Then get the aqlPath for this field from the WebTemplate.

Vorname:

```
Vorname-Path in CKM = at [openEHR-EHR-
COMPOSITION.person.v0]/content[openEHR-EHR-
ADMIN_ENTRY.person_data.v0]/data[at0001]/items[openEHR-EHR-
CLUSTER.person.v1 and name/value='Person']/items[openEHR-EHR-
CLUSTER.structured_name.v0 and name/value='Name']/items[at0002 and
name/value='Vorname']
```

```
AQLPath in WebTemplate =
/content[openEHR-EHR-
ADMIN_ENTRY.person_data.v0]/data[at0001]/items[openEHR-EHR-
CLUSTER.person.v1 and name/value='Person']/items[openEHR-EHR-
CLUSTER.structured_name.v0 and name/value='Name']/items[at0002 and
name/value='Vorname']/value
```

Nachname:

```
Nachname = at [openEHR-EHR-COMPOSITION.person.v0]/content[openEHR-
EHR-ADMIN_ENTRY.person_data.v0]/data[at0001]/items[openEHR-EHR-
CLUSTER.person.v1 and name/value='Person']/items[openEHR-EHR-
CLUSTER.structured_name.v0 and name/value='Name']/items[at0005 and
name/value='Familiennome']
```

```
AQLPath in WebTemplate =
/content[openEHR-EHR-
ADMIN_ENTRY.person_data.v0]/data[at0001]/items[openEHR-EHR-
CLUSTER.person.v1 and name/value='Person']/items[openEHR-EHR-
CLUSTER.structured_name.v0 and name/value='Name']/items[at0005 and
name/value='Familiennome']/value
```

- Remove the parts with the pattern "items[openEHR-EHR-CLUSTER.person.v1 and name/value='Person']" and substitute them with items[openEHR-EHR-CLUSTER.person.v1,'Person'] since this is how it works with the EHRBase
- Add a /value at the end of the AQL-Paths for vorname and nachname because the base-path gives the following element:

```
{
  "_type": "DV_TEXT",
  "value": "Hertmann"
}
```

- Resulting AQL-Query that need to be sent to `{{host}}/rest/openehr/v1/query/aql`

```
SELECT
  DISTINCT c/uid/value as compId,
  e/ehr_status/subject/external_ref/id/value as subjectId,
  c/content[openEHR-EHR-
ADMIN_ENTRY.person_data.v0]/data[at0001]/items[openEHR-EHR-
CLUSTER.person.v1,'Person']/items[openEHR-EHR-
CLUSTER.structured_name.v0,'Name']/items[at0002]/value/value as vorname,
  c/content[openEHR-EHR-
ADMIN_ENTRY.person_data.v0]/data[at0001]/items[openEHR-EHR-
CLUSTER.person.v1,'Person']/items[openEHR-EHR-
CLUSTER.structured_name.v0,'Name']/items[at0005]/value as nachname
FROM EHR e
CONTAINS COMPOSITION c
  CONTAINS ADMIN_ENTRY a
    CONTAINS CLUSTER[openEHR-EHR-CLUSTER.person.v1]
```

- Result:

```
{
  "q": "\n    SELECT \n          DISTINCT c/uid/value as compId,\n    e/ehr_status/subject/external_ref/id/value as subjectId,\n    c/content[openEHR-EHR-
ADMIN_ENTRY.person_data.v0]/data[at0001]/items[openEHR-EHR-
CLUSTER.person.v1,'Person']/items[openEHR-EHR-
CLUSTER.structured_name.v0,'Name']/items[at0002]/value/value as vorname,\n    c/content[openEHR-EHR-
ADMIN_ENTRY.person_data.v0]/data[at0001]/items[openEHR-EHR-
CLUSTER.person.v1,'Person']/items[openEHR-EHR-
CLUSTER.structured_name.v0,'Name']/items[at0005]/value/value as nachname\n    FROM EHR e\n    CONTAINS COMPOSITION c \n    CONTAINS ADMIN_ENTRY a\n    CONTAINS CLUSTER[openEHR-EHR-CLUSTER.person.v1]\n  ",
  "columns": [
    {
      "path": "/uid/value",
      "name": "compId"
```



```

    },
    {
      "path": "/ehr_status/subject/external_ref/id/value",
      "name": "subjectId"
    },
    {
      "path": "/content[openEHR-EHR-
ADMIN_ENTRY.person_data.v0]/data[at0001]/items[openEHR-EHR-
CLUSTER.person.v1,'Person']/items[openEHR-EHR-
CLUSTER.structured_name.v0,'Name']/items[at0002]/value/value",
      "name": "vorname"
    },
    {
      "path": "/content[openEHR-EHR-
ADMIN_ENTRY.person_data.v0]/data[at0001]/items[openEHR-EHR-
CLUSTER.person.v1,'Person']/items[openEHR-EHR-
CLUSTER.structured_name.v0,'Name']/items[at0005]/value/value",
      "name": "nachname"
    }
  ],
  "rows": [
    [
      "2ff4c284-a8b1-49e6-af0b-cf44a9a0b3ac",
      "89a21f4c-ed1b-4e0d-877a-4d03c5f772f6::local.ehrbase.org::1",
      "fakeid123",
      "Sigmar",
      "Tasche"
    ],
    [
      "489ee219-012a-404b-b5d3-470a8a9ed6bb",
      "aef91da6-ecb6-40a6-ae0-53a653358539::local.ehrbase.org::1",
      "fakeid125",
      "Sigmar",
      "Tasche"
    ],
    [
      "a80477cd-54d8-4552-a27d-2f3dd3d24ca9",
      "3d252a4e-7a6f-4895-8152-3efa4b5e399c::local.ehrbase.org::1",
      "fakeid124",
      "Rudi",
      "Klapp"
    ],
    [
      "b514ab79-78f9-4988-a942-37f733a16590",
      "088a55d2-1e83-4111-9f7e-229f423051f2::local.ehrbase.org::1",
      "fakeid127",
      "Ferdinand",
      "Hertmann"
    ],
    [
      "b7aada04-2b38-4ec7-9c81-961b0565e4ee",
      "73180ccb-3065-49fe-b9f8-ec39a4e1e7cc::local.ehrbase.org::1",
      "fakeid126",
      "Kristin",

```

```
        "Löffler"  
      ]  
    ]  
  }
```

- The result holds all the information that we queried:

```
EHR-ID: "b7aada04-2b38-4ec7-9c81-961b0565e4ee",  
Composition-ID: "73180ccb-3065-49fe-b9f8-ec39a4e1e7cc::local.ehrbase.org::1",  
Subject-ID: "fakeid126",  
Vorname: "Kristin",  
Familiennamen: "Löffler"
```