



Development of a Web Application for the Comparison of Medical Diagnoses

Appendix: User Guide

| | |
|------------------|------------------------------|
| Course of study: | Medical informatics |
| Author: | Kündig Moritz, Tschanz Roger |
| Supervisor: | Thomas Bürkle |
| Date: | 21.01.2021 |

Inhaltsverzeichnis

| | | |
|-------|--|----|
| 1 | Introduction | 2 |
| 2 | User Guide | 3 |
| 2.1 | Start Page / Login-Window / Registration | 3 |
| 2.2 | Dashboard | 4 |
| 2.3 | Navigation | 4 |
| 2.4 | Operate Datasets | 5 |
| 2.4.1 | Add a Dataset | 5 |
| 2.4.2 | Detailed Dataset View | 7 |
| 2.5 | Operate Analysis | 8 |
| 2.5.1 | Run Analysis | 8 |
| 2.5.2 | Run Analysis | 10 |
| 2.5.3 | Download Results | 10 |
| 2.6 | Operate Classifications | 11 |
| 2.6.1 | Prepare Classifications | 11 |
| 2.6.2 | Add a Classification | 14 |
| 2.7 | Operate Code Lookup | 14 |
| 2.8 | Operate Parser | 15 |
| 2.9 | Imprint und Privacy | 16 |
| 3 | Literature | 17 |

1 Introduction

This document is part of the bachelor thesis of the students Roger Tschanz and Moritz Kündig of the Bern University of Applied Sciences BFH. It describes the use of the web application "Armin2.0" developed during the project. The web application measures the distances between single and set diagnosis using several algorithms, divided into the levels "information content (IC)", "code-level similarity (CLS)" and "set-level similarity (SLS)". The supported classifications are ICD-9 WHO, ICD-10 WHO and ICD-10 GM.

Armin has been designed and tested for use with Google Chrome (version: 87.0.4280.141) and Microsoft Edge (version: 87.0.664.75) on Windows 10. The application also runs in other browsers such as Firefox and Brave and on other operating systems such as MacOS and Ubuntu, but the latter are not officially supported. In case certain functionalities like uploading do not work, switch to Windows 10 and the respective Google Chrome or MS Edge versions.

The application is available at the following URL: www.right-icd.de

2 User Guide

2.1 Start Page / Login-Window / Registration

The start page appears in the following format. If you are no user yet, create a new user via the “Register” field. Then click “Register” to confirm the process. To register, you need your first and last name, a unique email address, a password of at least 8 characters and the name of your medical institution (see Figure 1:).

Register

| | |
|------------------|--|
| First Name | <input type="text" value="Hans"/> |
| Name | <input type="text" value="Muster"/> |
| Institute | <input type="text" value="Berner Fachhochschule"/> |
| Country | <input type="text" value="Switzerland"/> |
| E-Mail Address | <input type="text" value="hans.muster@bfh.ch"/> |
| Password | <input type="password" value="....."/> |
| Confirm Password | <input type="password" value="....."/> |

[Privacy Statement & Impressum](#)

Register

Figure 1: Registration Screen

If you are already registered, please login by entering your email address and your password on the login page (see Figure 2: Login).

Figure 2: Login Screen

Alternatively, you can also log in via your existing personal Google account. In this case, you are redirected to the Google login page.

2.2 Dashboard

After a successful login, you are redirected to the dashboard. This page lists all personal datasets on the left side and the available classifications on the right side. In Figure 3, a dataset has already been created and the user has access to a classification (see Figure 3: Dashboard).

Figure 3: Dashboard Screen

2.3 Navigation

Armin2.0 can be navigated by using the menu in the header (i.e. on top of the page) and in the footer (i.e. at the bottom of the page). The main menu in the header has the following five options (see Figure 4: Menu View in the Header):

1. "Datasets" to manage and run datasets.
2. "Classifications" to add new classifications, e.g. a new ICD-10 version, or to examine existing classifications.
3. "Code Lookup" to search the uploaded classifications.
4. "Parser" to parse free text diagnoses into ICD codes.



Figure 4: Menu View in the Header

You can also access the menu items via the footer. In addition, you find the imprint and the privacy policy here (see Figure 5: Menu View in the Footer).

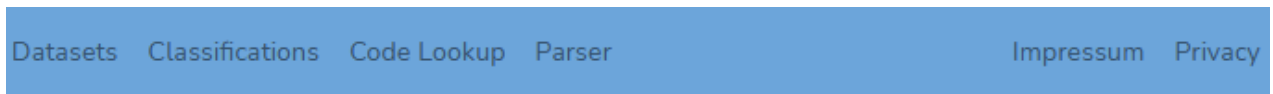


Figure 5: Menu View in the Footer

2.4 Operate Datasets

Access the "Datasets" module via "Datasets". The module enables the upload of datasets containing diagnoses to be compared. After uploading, you can run any analyses with a dataset.

2.4.1 Add a Dataset

After clicking "Add Dataset", the form for uploading a new dataset opens (see Figure 6: Form View for Adding).

The analysis name and the classification are required fields. All ICD classifications which have codes and have been generated in this application are available under "Classifications". In Figure 6, the *ICD-10 WHO Version 2016* is selected. Inserting a description of the dataset is optional.

You can then add a CSV file including the diagnoses by moving it into the "Drag and drop file here" field. The file is validated, marked by a green field if successful, and can then be configured (see Figure 7: Upload View for).

If the uploaded CSV file has headers, please select the field "Has headers" (see the red frame in Figure 7). Then map the columns from the CSV file manually.

Figure 6: Form View for Adding Datasets

Strictly observe the meaning of the fields in the following mapping

| Column | Application Name | Description | Optional | Comment |
|--------|------------------|----------------------------|----------|--|
| 1 | Patient ID | The patient-identification | No | Can be any patient ID as an integer, e.g. 1, 2, 3, 4, etc. |
| 2 | Code | The ICD-code | No | The ICD code must be present in the selected classification of the dataset, in this case 'ICD-10 WHO 2016'. Otherwise, the calculation cannot be made. |
| 3 | Set ID | The set-ID | No | Indicates to which set the ICD code in column 2 (above) belongs to. This can be presented as any integer, e.g. 1,2,3,1,3, etc. |
| 4 | Code main | Main diagnosis | No | Indicates whether the ICD-code is a principal diagnosis. Value 1 = principal diagnosis, value 0 = no principal diagnosis |
| 5 | Code current | Current diagnosis | Yes | Indicates whether the ICD code is a current diagnosis to be taken into account in the calculation. Value 1 = current diagnosis, value 0 = no current diagnosis |
| 6 | Code suspected | Suspected diagnosis | Yes | Indicates whether the ICD code is a suspected diagnosis to be taken into account in the calculation. Value 1 = suspected diagnosis, value 0 = no suspected diagnosis |
| 7 | Code Excluded | Exclude diagnosis | Yes | Indicates whether the ICD code is a current diagnosis to be taken into account in the calculation. Value 1 = exclude diagnosis, value 0 = include diagnosis. |

Table 1: Structure of a Diagnosis Set

The following list shows examples of valid and invalid CSV file formats.

✓ **Valid format**

Semicolon separation is allowed

```
patid;code;set;main;current;suspected;e
123;S05.0;1;1;1;0;0
123;S05.0;2;1;1;0;0
abc;Z72.0;1;1;1;0;0
abc;F10.1;1;0;1;0;0
abc;Z72.0.9;2;0;1;0;1
```

✓ **Valid format**

No column labelling required

```
123;S05.0;1;1;1;0;0
123;S05.0;2;1;1;0;0
abc;Z72.0;1;1;1;0;0
abc;F10.1;1;0;1;0;0
abc;Z72.0.9;2;0;1;0;1
```

✓ **Valid format**

Additional columns are allowed but not recommended

```
patid;code;set;main;current;suspected;excluded;name
123;S05.0;1;1;1;0;0;Max
123;S05.0;2;1;1;0;0;Max
abc;Z72.0;1;1;1;0;0;Marta
abc;F10.1;1;0;1;0;0;Marta
abc;Z72.0.9;2;0;1;0;1;Marta
```

✓ **Valid format**

Only the obligatory four columns are required

```
patid;code;set;main
123;S05.0;1;1
123;S05.0;2;1
abc;Z72.0;1;1
abc;F10.1;1;0
abc;Z72.0.9;2;1
```

New Dataset

Upload patients and diagnoses to dataset

Dataset Details

Description
New Dataset Description

Classification
ICD-10 WHO 2016

Last Updated
1 second ago

Delete

Note that no personal patient data will be uploaded. They are not needed for the calculation. Any liability is excluded. [Read more](#)

Configuration

☒ Has headers

| Field | CSV Column |
|----------------|------------|
| Patienten ID * | patid |
| Code * | code |
| Set ID * | set |
| Code Main * | main |
| Code Current | current |
| Code Suspected | suspected |
| Code Excluded | excluded |

Cancel Create

Figure 7: Upload View for Datasets

Important note: Only upload completely anonymised files, which do not allow any conclusions to be drawn about the patients. A similar note is also displayed above the “Drag and drop file here” field.

2.4.2 Detailed Dataset View

After creating the dataset via “Create”, you are redirected to the detailed view. You can also select and display the detailed view via the “Dataset” item in the main menu. In this view, you can also display the individual and set diagnoses of the respective patients via “Show Patients” (see Figure 8: Detailed View of a Dataset). In addition, the presentation of the different “diagnosis types” is explained.

By clicking “Start analysis”, the analysis is created (see red frame Figure 8). Another form opens before the analysis is run (see Figure 9: Form View for s).

New Dataset

Analysis overview for dataset: New Dataset

Start an analysis

Dataset Details

Description

New Dataset Description

Classification

ICD-10 WHO 2016

Last Updated

4 minutes ago

Delete

Patients

Show Patients

Info: Main diagnoses are marked in **bold**, suspected diagnoses are in brackets () and excluded diagnoses are crossed-out.

| Paitent ID | First Diagnoses | Second Diagnoses |
|------------|--|-----------------------------------|
| 1 | S050 | S050 |
| 2 | Z720 F101 I469 H570 | Z720 F101 J111 |
| 3 | Z720 | F101 Z720 |
| 4 | H570 J111 R58 (F31) | H570 J111 R58 F32 |

Figure 8: Detailed View of a Dataset

2.5 Operate Analysis

Any number of analyses can be run based on one dataset. On the other hand, an analysis is always based on only one uploaded data set including its classification. If, for example, a dataset is to be compared with another classification version, a new dataset must be created to run this analysis.

2.5.1 Run Analysis

The following information is required to create a new analysis:

1. The analysis name
2. An analysis description is optional.
3. The algorithms to be used for comparison. In below example, the algorithms are "Level" at IC level, "Wu and Palmer" at CLS level and "Jia et al." at SLS level.

Dataset Details

Description

New Dataset Description

Classification

ICD-10 WHO 2016

Last Updated

9 minutes ago

Delete

Analysis Name *

New Analysis

Analysis Description

New Analysis Description

Information Content (IC) *

Level

$levels(a \rightarrow r)$

Concept-level similarity (CS) *

Wu & Palmer

$1 - \frac{2IC(c)}{IC(a) + IC(b)}$

Set-level similarity (SS) *

Jia et al.

$\frac{1}{|A| \cdot |B|} \cdot \sum_{a \in A, b \in B} CS(a, b)$

Cancel

Create

Figure 9: Form View for Adding Analyses

| No. | Algorithm and Comments |
|----------------------------|--|
| IC.01 level | <p>The level of the ICD tree is taken as the value. Here: 0 is Root, i.e. ICD-10 GM and 5 are the lowest level This formula fits better than IC.02 to separate different concepts more clearly under the same chapter [1].</p> |
| IC.02 ontology-based score | <p>Definitions:</p> <ul style="list-style-type: none"> Subsumer: All parent nodes above up to and including the root level. The leaves(r) are all leaves or codes of an ICD tree, the leaves(a) are all leaves below of the code to be determined. <p>This formula has a better selectivity at chapter level than the level-based approach of the IC.01 formula [1].</p> |
| CLS.01 binary | <p>The binary formula gives the result 0 or 1.</p> <ul style="list-style-type: none"> 0 = the diagnosis are equal 1 = the diagnoses are unequal |
| CLS.02 Wu and Palmer | <p>Definitions:</p> <ul style="list-style-type: none"> Variable c = least common superconcept (LCA) The numerator determines where the formulas “come together”. <p>Context:</p> <ul style="list-style-type: none"> 0 = the diagnoses are close together or equal 1 = the diagnoses are far apart or unequal |
| CLS.03 Li et al. | <p>Context: Inverse to Wu and Palmer and binary</p> <ul style="list-style-type: none"> 0 = the diagnoses are far apart or unequal 1 = the diagnoses are close together or equal <p>Alpha and beta are constants and are used for scaling. They are set to alpha = 0.2 und beta = 0.6 in this implementation, according to Jia et al. [2].</p> |
| CLS.04 Steps | Counts the steps between the first and second diagnosis by using the LCA. |
| SLS.01 Dice | <p>This formula calculates the arithmetic mean of sets A and B. $A \cap B$ contains all diagnoses in Set A and B</p> <p>Context:</p> <ul style="list-style-type: none"> 0 = close together or the sets are identical 1 = far apart or the sets are unequal. |
| SLS.02 Jaccard | Context: |

| | |
|-------------------|--|
| | <ul style="list-style-type: none"> 0 = close together or the sets are identical. 1 = far apart or the sets are unequal. |
| SLS.03 Cosine | Similar formula as SLS.01, but calculates the geometric mean. In theory, the geometric min. is equally exact or more exact than the arithmetic mean. |
| SLS.04 Overlap | Calculates the cardinality of the intersection of both sets A and B by the cardinality of the smaller set. |
| SLS.05 Jia et al. | The double sum calculates the distances of each diagnosis a from Set A with each diagnosis b from Set B. The average is then calculated from this sum. |

Table 2: Implemented Algorithms Including Comments [1, 3]

2.5.2 Run Analysis

Click Create to run the analysis and to calculate the results of the previously selected algorithms. The "Analysis Details" field lists which algorithms were used and offers the possibility to download the results. The field "Patient Results" shows a section of the results and, if a calculation could not be performed, a comment.

Note: Depending on the size of the CSV files, this step can take a few minutes.

Analysis Details

Classification

ICD-10 WHO Version 2019

Patient Count

4

Information Content (IC)

Level

Concept-level similarity (CS)

Wu & Palmer

Set-level similarity (SS)

Jia et al.

Download Result

Delete

Patient Results (Excerpt: first 100 records)

| Internal Pat ID | Value | Remark |
|--------------------------------------|-------------------|--------|
| 5659d45e-25b5-11eb-adc1-0242ac120002 | 0 | |
| 00e265a8-25b6-11eb-adc1-0242ac120002 | 0.777777777777778 | |
| 0851c8f6-25b6-11eb-adc1-0242ac120002 | 1 | |
| 42288ba4-8a5a-43d6-b514-afdfd314c32d | 0.777777777777778 | |

2.5.3 Download Results

The results can be downloaded as a ZIP archive which contains 4 files:

1. The results of the IC hierarchy (*.ci.csv)
2. The CLS results (*.cls.csv)
The similarities of the individual diagnoses are compared here. All individual diagnoses of the first set are compared with all individual diagnoses of the second set in the format of a cartesian product.
3. The SLS results (*.sls.csv)
The results of the compared diagnosis sets are displayed here, congruent with the "Patient Results" field in the application.
4. A readme-file (Read me.txt)
General information and metadata from the analysis are listed here, e.g. the analysis date and the analysis details.

2.6 Operate Classifications

Classifications can be uploaded in this module. You need them to determine on which classification the diagnoses to be compared are based.

2.6.1 Prepare Classifications

Currently, any user can create a classification. You need a CSV with the three columns: Code, Code Name and the respective Parent. The process depicted below is specific to the ICD-10 GM 2018 classification and may differ if you obtain the data from a different source or if you follow the process using a different classification. It is important in any case that you follow the principles in Table 2: Structure of the Classification Data for Upload:

| Column | Application Name | Optional | Description and Comment |
|--------|------------------|----------|--|
| 1 | Code | No | The respective ICD code, e.g. A41. This code must be unique. |
| 2 | Code name | No | A corresponding name for the code, e.g. "Other sepsis". |
| 3 | Parent | No | The parent object is empty if the respective code is a chapter, i.e. it is on the top level. Otherwise, the parent indicates the direct parent node, e.g. "A30-A49" for other bacterial diseases. A reference to a parent node is only possible if the corresponding 'code' already exists in the database. |

Table 2: Structure of the Classification Data for Upload

You can access the data for ICD-10 GM 2018 here:

<https://www.dimdi.de/dynamic/downloads/klassifikationen/icd-10-gm/vorgaenger/icd10gm2018.zip>

To view the data, extract the zip file after downloading and search it for the systematic directory including the CSV files. According to the readme naming convention you recognise the file as follows:

| 1. bis 3. Stelle | Klassifikation u. Band |
|------------------|--------------------------------|
| x1g | ICD-10-GM Systematik |
| x3g | ICD-10-GM Alphabet |
| x1w | ICD-10-WHO Systematik (Band 1) |
| x2w | ICD-10-WHO Regelwerk (Band 2) |
| x3w | ICD-10-WHO Alphabet (Band 3) |
| p1s | OPS Systematik |
| p2s | OPS Alphabet |

| 4. Stelle | Fassung |
|-----------|----------------------|
| a | Aktualisierungsliste |
| b | Buchfassung |
| e | EDV-Fassung |
| m | Metadaten |
| r | Regelwerk |
| u | Überleitung |

| 5. Stelle | Format |
|-----------|------------|
| h | HTML |
| o | ODT |
| p | PDF |
| r | RTF |
| t | TEXT (CSV) |
| c | ClaML/XML |

Figure 10: Extract from the File Index from the ICD-10 GM 2018 Download

From the x1gmt folder you need the data of the chapters, groups and codes:

| ls > icd10gm2018 > x1gmt2018 > Klassifikationsdateien | | | | |
|---|------------------|---------------|----------|--|
| <input type="checkbox"/> Name | Date modified | Type | Size | |
| <input checked="" type="checkbox"/> icd10gm2018syst_gruppen.txt | 18/01/2021 21:22 | Text Document | 16 KB | |
| <input checked="" type="checkbox"/> icd10gm2018syst_kapitel.txt | 18/01/2021 21:22 | Text Document | 2 KB | |
| <input checked="" type="checkbox"/> icd10gm2018syst_kodes.txt | 18/01/2021 21:22 | Text Document | 3.908 KB | |
| <input type="checkbox"/> icd10gm2018syst_kodes.txt | 18/01/2021 21:22 | Text Document | 3.908 KB | |

Figure 11: Chapter, Groups and Classification Files

Transfer the three tables into three sheets in Excel. To do that, use the "Import data from Text/CSV" function under the "Data" tab (see Figure 12: Selecting the "Import from Text/CSV" Function in Excel). When importing you must set the separator according to the source file. Then set the correct character set to prevent errors with umlauts. (see Figure 13: Extract from the Import Function with Selected Parameters).

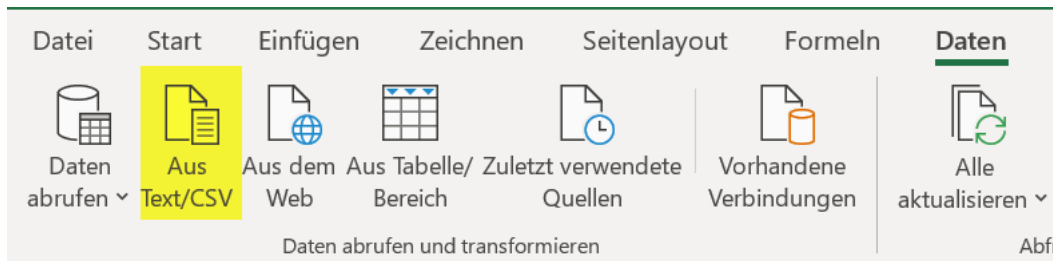


Figure 12: Selecting the "Import from Text/CSV" Function in Excel

| | | |
|---|---------------------------|--|
| Dateiursprung 65001: Unicode (UTF-8) | Trennzeichen Semikolon | Datentyperkennung Basierend auf den ersten 200 Zeilen |
|---|---------------------------|--|

| Column1 | Column2 |
|---------|---|
| 1 | Bestimmte infektiöse und parasitäre Krankheiten |
| 2 | Neubildungen |

Figure 13: Extract from the Import Function with Selected Parameters

After the import, the sheets should look as follows. Their names have been changed for easier understanding:

| | | |
|---------|---------|----------|
| 4 T | X | 1 A00 |
| 1 T | Y | 1 A00 |
| Kapitel | Gruppen | Codes |
| | | Tabelle1 |

Figure 14: Extract of the Excel sheets after the import

Excel automatically sets headers because none were set in the file. The headers of the individual files are listed in the readme file of the x1gmt folder.

Now merge all three sheets into a new one as described in the following steps. Start with the chapters, then add the groups and finally add the codes. Strictly keep to this order when adding the codes so that the chapters and groups are imported first and there is no error within the application. It is also important that each code is unique per classification.

Chapters: Copy the chapters as they are. Leave the column "Parent" empty:

| A | B | C | D | E |
|------|---|--------|---|---|
| Code | Name | Parent | | |
| 1 | Bestimmte infektiöse und parasitäre Krankheiten | | | |
| 2 | Neubildungen | | | |
| 3 | Krankheiten des Blutes und der blutbildenden Organe s | | | |
| 4 | Endokrine, Ernährungs- und Stoffwechselkrankheiten | | | |

Figure 15: Extract of the New Sheet after Transferring the Classifications

Groups: The sheet with the groups consists of the following fields:

- Column 1: First three-digit of the group, 3 characters
- Column 2: Last three-digit of the group, 3 characters
- Column 3: Chapter number, max. 2 characters
- Column 4: Group title, up to 255 characters

Edit the four columns as follows to obtain three columns:

- New column 1: Combine columns 1 and 2 from above by using a hyphen, e.g. 'A00-A09'.
- New column 2: Use the group title from column 4, i.e. "Infectious intestinal diseases".
- New column 3: Use column 3, i.e. the chapter number.

Then attach the three columns to the chapter data from above. The chapter number of the third column should point to the respective parent chapter above:

| | A | B | C | D |
|----|---------|---|--------|---|
| 1 | Code | Name | Parent | |
| 2 | 1 | Bestimmte infektiöse und parasitäre Krankheiten | | |
| 3 | 2 | Neubildungen | | |
| 4 | 3 | Krankheiten des Blutes und der blutbildenden Organe sowie bestimmte Störungen mit Bet | | |
| 5 | 4 | Endokrine, Ernährungs- und Stoffwechselkrankheiten | | |
| 22 | 21 | Faktoren, die den Gesundheitszustand beeinflussen und zur Inanspruchnahme des Gesund | | |
| 23 | 22 | Schlüsselnummern für besondere Zwecke | | |
| 24 | A00-A09 | Infektiöse Darmkrankheiten | 1 | |
| 25 | A15-A19 | Tuberkulose | 1 | |
| 26 | A20-A28 | Bestimmte bakterielle Zoonosen | 1 | |
| 27 | A30-A49 | Sonstige bakterielle Krankheiten | 1 | |
| 28 | A50-A64 | Infektionen, die vorwiegend durch Geschlechtsverkehr übertragen | 1 | |

Figure 16: Extract of the New Spreadsheet after Transferring the Groups

Codes: Note that the sheet with the codes contains all codes from level 3 to 5. To create the new sheet, you need the following columns:

- Column 1: Classification level, 1 character
- Column 4: Chapter number, max. 2 characters
- Column 8: Key number without period, dash, asterisk and exclamation mark, up to 5 characters
- Column 9: Class title composed of components of the titles of the three-digit, four-digit and five-digit codes, if any, up to 255 characters.

Edit the four columns as follows to obtain three columns:

- New column 1: Use column 8 as is.
- New column 2: Use column 9.
- New column 3: If the code in column 1 is in classification level 3, create a VLOOKUP on the group table to get to the composite group code, e.g. 'A00-A09'. If the classification level is greater than 3, use column 8 and delete the last character.

The final result must look as follows:

| | | | | |
|-----|---------|--|---------|--|
| 264 | U99-U99 | Nicht belegte Schlüsselnummern | 22 | |
| 265 | A00 | Cholera | A00-A09 | |
| 266 | A000 | Cholera durch Vibrio cholerae O:1, Biovar cholerae | A00 | |
| 267 | A001 | Cholera durch Vibrio cholerae O:1, Biovar eltor | A00 | |
| 268 | A009 | Cholera, nicht näher bezeichnet | A00 | |
| 269 | A01 | Typhus abdominalis und Paratyphus | A00-A09 | |
| 270 | A010 | Typhus abdominalis | A01 | |
| 271 | A011 | Paratyphus A | A01 | |
| 272 | A012 | Paratyphus B | A01 | |

Figure 17: Extract of the new spreadsheet after transferring the codes

Save the new sheet as a CSV file to get it ready for upload.

2.6.2 Add a Classification

After clicking on "Add classification", the form for uploading a new classification opens.

You need a name and the language; currently English and German are available. Adding a description is optional.

Then upload the CSV file prepared in chapter 2.8.1 Prepare Classifications via the "Drag and drop file here" field. The file is validated, marked with a green field if successful and can then be configured.

If the uploaded CSV file has headers, select the field "Has headers".

Then manually select the columns containing the ICD code, the concept name and the parent node. You can only continue when you have selected three different columns. Complete the classification by clicking on "Create".

Note: The upload and allocation can take a few seconds to several minutes depending on the size of the file.

Figure 18: Upload Classification Part 2

2.7 Operate Code Lookup

This module allows to search for ICD codes or names in the ICD classifications that have been added to "Armin".

Note: You can also use the Code Lookup via the 'Classifications' module.

First select the classification in the field on the left, see red frame in Figure 19: Code-Lookup .

Figure 19: Code-Lookup View

You can then search for an ICD code, a disease name or a keyword in the field on the right. The following formats are supported:

- A01: A search of a three-digit code returns all 3, 4 and 5-digit codes A01, i.e. A01, A010, A011, A012 etc. in the example below.
- A010: A search of a four-digit code returns all 3-digit and 4-digit codes.
- The keyword search returns all codes that contain this word in any form. For example, “typhoid” returns codes with “paratyphoid” and “typhus abdominalis”.
- The keyword search for “paratyphoid”, on the other hand, only returns codes with this keyword.

In the following example, “A01” is searched for (see Figure 20: View Search Results under “Code Lookup”). The columns “Level” and “Ontology-based Score” show the corresponding IC of the code.

| ICD-10 WHO 2016 | ▼ | A01 | Search |
|-----------------|--------------------------------|-------|----------------------|
| Code # | Name | Level | Ontology-based Score |
| A01 | Typhoid and paratyphoid fevers | 3 | 3.673451 |
| A010 | Typhoid fever | 4 | 3.946452 |
| A011 | Paratyphoid fever A | 4 | 3.946452 |
| A012 | Paratyphoid fever B | 4 | 3.946452 |
| A013 | Paratyphoid fever C | 4 | 3.946452 |
| A014 | Paratyphoid fever, unspecified | 4 | 3.946452 |

Figure 20: View Search Results under “Code Lookup”

Note: The text search is not case sensitive.

2.8 Operate Parser

This module enables the parsing of free text diagnoses into ICD codes.

Important to note:

- This parser is simply an interface that sends the entered diagnostic texts to an API of an existing parser of the UNZ University Emergency Centre; there the texts are parsed and then the ICD codes are sent back to “Armin application”. The API is available (as of December 2020) at <https://www.covidtrack.ch/parser>.
- The authors of the Armin2.0 application had no influence on the functioning and internal processes of the parser. For further questions, please use the contact information provided in the imprint.
- The parser currently only supports the parsing of German diagnosis texts of the ICD classification ICD-10 GM.

The result contains the following 6 columns. The columns are defined similarly to those in Table 1: Structure of a Diagnosis Set.

1. ICD-10 code
2. Code-text
3. Selection of whether the diagnosis is a principal or secondary diagnosis
4. Display whether the diagnosis is up to date.
5. Display whether the diagnosis is a suspected diagnosis.
6. Display whether the diagnosis can be omitted.

Usually activated when the free text contains a negex such as “none” or “without”.

Parser

Enter a German free text diagnosis and the respective diagnoses are parsed

Diagnosis

1.1.2021 - Unklare ziehende rezidivierende Kopfschmerzen, kein Husten
1.1.2020 - Infekt der oberen Atemwege, Vd. a. Migräne mit visueller Aura

[Cancel](#)
[Parse](#)

[Download](#)

| Code | Code Text | Main | Current | Suspected | Excluded |
|-------|----------------------------|------|---------|-----------|----------|
| R51 | Kopfschmerzen | ✓ | ✓ | ✗ | ✗ |
| J06.9 | Infekt der oberen Atemwege | ✗ | ✓ | ✗ | ✗ |
| G43 | Migräne | ✗ | ✓ | ✗ | ✗ |
| R05 | Husten | ✗ | ✓ | ✗ | ✓ |

Figure 21: Diagnosis-Parsing Screen

2.9 Imprint und Privacy

These menu items contain general information about the operator of this website, the “Terms and Conditions” and the setting of privacy options.

3 Literature

References

1. Jia Z, Lu X, Duan H et al. Using the distance between sets of hierarchical taxonomic clinical concepts to measure patient similarity. *BMC Med Inform Decis Mak* 2019;19(1):91.
2. Jia Z, Zeng X, Duan H et al. A patient-similarity-based model for diagnostic prediction. *International Journal of Medical Informatics* 2020;135:104073.
3. Girardi D, Wartner S, Halmerbauer G et al. Using concept hierarchies to improve calculation of patient similarity. *J Biomed Inform* 2016;63:66–73.