Mapping of German Language Free Text Laboratory Results to SNOMED CT

Andrea Riedel, Noemi Deppenwiese, Simon Winkler |
University Hospital Erlangen-Nuremberg

Introduction
Methods
Results
Discussion

Abstract

Even though laboratories are highly digitalized, free text entries are still part of the daily laboratory routine to save laboratory results. Especially for retrospective and current analysis free text entries will be a problem for statistical analysis. The goal of the project is the mapping of German free text entries in laboratory results to the English-language terminology SNOMED CT to enhance semantic and syntactic interoperability.[1] SNOMED CT is used in the different laboratories of the University Hospital Erlangen-Nuremberg to transfer abbreviations and colloquial language from the free text entries into standardized technical language. This project could be a model for research networks over multiple German hospital sites and therefore shows challenges and chances of such an endeavor.

Learning Objectives

- Learn the importance of structured laboratory data for local and international data analysis for research and patient treatment
- Understand the challenges of the combination of similar free text results of different laboratories of one hospital
- Get an insight into the complex representation of different detail levels of laboratory information
- Understand the complexity of translating German language free text entries into a fix English terminology

Background

- Medical Informatics Initiative (MII) established data integration centers (DIC) at university medical sites [2]
- DICs are facilities for the data selection from healthcare and research of German university hospitals
- Uses international standards to enable semantical interoperability [3]
- Structured data should be usable for national and international medical research and data analysis to improve patient treatment [4]

Scope

- German language free text entries of laboratory results
- Site: University Hospital Erlangen-Nuremberg [5]
- 17 individual laboratories, each with their own internal analyte codes
- LOINC codes have been used to map analytes since 2000
- Laboratory values (result of the measurement) still mainly free text
- Every laboratory has its own internal rules or abbreviation to shortly describe specific results
- Details and documentation is still highly depended on the laboratory staff
- Focus on German Corona Consensus Dataset (GECCO) analytes and most common analytes of the local laboratories [6]
- Could be a model for research networks over multiple German hospital sites

Mapping of German Language Free Text Laboratory Results to SNOMED CT

Andrea Riedel, Noemi Deppenwiese, Simon Winkler |
University Hospital Erlangen-Nuremberg

Introduction
Methods
Results
Discussion

Initial Situation

- Same situation may be communicated in different terms by different staff members or even the same staff member
- Terms may also differ between different laboratories, especially in case of errors
- All terms representing the same concept must be mapped to one SNOMED CT code
- Generally mapping to qualifier values is preferred because all other information can be inferred from context

Example

A sample may be *geronnen (clotted)*. In this case, laboratory staff may enter the full term, one of several abbreviations like *ger* or *ger.*, more than one word like *Blut geronnen* (*Blood clotted*), colloquial language like *leider geronnen* (*unfortunately clotted*) or even a sentence like *Blut war geronnen* (*Blood is clotted*). All terms are mapped to 281311008 |Clotted (qualifier value)| (and not 281279002 |Sample clotted (finding)|).

Challenges

- Laboratory-internal shortcuts for concepts may differ or collide between laboratories
- Mapping requires using the original (laboratory-specific) analyte code (in addition to the standardized LOINC code) to allow differentiating between laboratories and generate a correct value mapping
- In contrast, identical test result values may map to the same SNOMED CT code, even if they appear with distinct laboratory codes that map to distinct LOINCs

Example

A laboratory result value of “-” may indicate an error like a patient mix-up, in which case it needs to be mapped to 281267002 |Incorrect sample received (finding)|, or that no test result is available (373121007 |Test not done (qualifier value)|), but in another laboratory it can also be a valid result value corresponding to 260385009 |Negative (qualifier value)|.

Mapping of German Language Free Text Laboratory Results to SNOMED CT

Andrea Riedel, Noemi Deppenwiese, Simon Winkler |
University Hospital Erlangen-Nuremberg

https://drive.google.com/file/d/1Frlfce6-Z84aJJksgPHNG7Y_TTEWlH3S/view?usp=sharing

Mapping to one identical SNOMED CT code

Internal laboratory codes are mapped to LOINC first (orange), then the values are mapped to SNOMED CT (blue). The video shows how different laboratories have different internal laboratory codes for the same LOINC code. The differences are highlighted with red borders. Moreover, one internal laboratory code and the LOINC code can be connected to different values which express the same content. Nevertheless, the value is mapped to the same SNOMED CT code. Internal background information from the original laboratory code may be lost when only using LOINC.

https://drive.google.com/file/d/1ntMKqa0xLtEHjX_Q1soHVDQ-a-mP1nDs/view?usp=sharing

Mapping of one identical free text value to different SNOMED CT codes

This video shows how different laboratories have different internal laboratory codes with the same free text value. Nevertheless, the same free text value expresses different concepts and therefore is connected to different SNOMED CT codes. Again, internal background information from the original laboratory code may be lost when only using LOINC. This also applies to different numbers of “–” symbols in one value field.

Mapping of German Language Free Text Laboratory Results to SNOMED CT

Andrea Riedel, Noemi Deppenwiese, Simon Winkler |
University Hospital Erlangen-Nuremberg

Conclusions

- German language laboratory text values can be mapped to SNOMED CT
- Free text laboratory values may differ between staff members and laboratories
- Therefore also taking the analyte code into account is essential
- Once all free text values are mapped to SNOMED CT, inside knowledge is no longer necessary to interpret the laboratory results
- Analyses across multiple laboratories become easier because laboratory values for the same result are represented with the same SNOMED CT code

Future Directions

- Laboratory data integration processes need to include the terminology SNOMED CT as well as locally defined mapping rules
- Regular updates to the new SNOMED CT versions mean regular updates of the mappings
- Re-validation of the implemented mapping by laboratory specialist will be a long-term task
- Recent SNOMED International membership of Germany means a lot of initial regional activities are in progress and the German Translation Group (GTG) has started its activities in 2020
- In the future, a national extension will be essential for German specialties in the laboratory domain

snomedexpo.org
SNOMED CT EXPO 2022
Sept 29-30, 2022 × Lisbon, Portugal

Discussion

References

1. Thun,S. et al. (2017) Syntaktische und semantische Interoperabilität. In: Müller-Mielitz S., Lux T. (eds) E-Health-Ökonomie. Springer Gabler, Wiesbaden. https://doi.org/10.1007/978-3-658-10788-8_34
2. Semler,S.C. et al. (2018). German Medical Informatics Initiative. Methods of information in medicine, 57(S 01), e50–e56. https://doi.org/10.3414/ME18-03-0003
3. Medical Informatics Initiative: SNOMED CT: Häufig gestellte Fragen, at https://www.medizininformatik-initiative.de/de/snomed-ct-haeufig-gestellte-fragen (last accessed 12.03.2022)
4. Medical Informatics Initiative: The Medical Informatics Initiative's core data set, at https://www.medizininformatik-initiative.de/en/medical-informatics-initiatives-core-data-set (last accessed 12.08.2022)
5. Prokosch,H.U. et al. (2018). MIRACUM: Medical Informatics in Research and Care in University Medicine. Methods of information in medicine, 57(S 01), e82–e91. https://doi.org/10.3414/ME17-02-0025
6. Sass,J. et al. (2020). The German Corona Consensus Dataset (GECCO): a standardized dataset for COVID-19 research in university medicine and beyond. BMC medical informatics and decision making, 20(1), 341. https://doi.org/10.1186/s12911-020-01374-w