

Graph-based creation of IDMP-conformant FHIR Medication instances using Neo4j: An evaluation study

Erzeugung IDMP-konformer FHIR Medication-Objekte unter Einsatz einer Neo4j-Graphdatenbank: Eine Evaluationsstudie

Abstract

Introduction: The availability of syntactically and semantically interoperable data is a critical prerequisite to high-quality research in the medical field. For medication data, the FHIR Core Data Set of the Medical Informatics Initiative Germany provides a powerful data model based on the IDMP standards that supports high levels of structure and quality. However, populating this model requires detailed medication information, which is often not available at the data integration centers.

Objective: To address this issue, this study aims to implement a software to generate MII CDS-compliant FHIR *Medication* objects based on pharmaceutical databases and to assess their conformity to standards.

Methods: Medication data from the MMI Pharmindex was loaded into a Neo4j graph database. Both manual and automatic mapping techniques were used to enrich the data with various terminologies. Finally, the resulting graph contents were exported into FHIR instances.

Results: The export generated approximately 70,000 FHIR *Medication* instances conforming to the MII Core Data Set. The inclusion of terminologies and IDMP compliance were partially successful, with missing data in the source datasets proving to be the primary issue.

Conclusion: Our results highlight the need for higher-quality data on pharmaceuticals available in Germany. Nevertheless, the generated Medication instances can accelerate the provision of a complete medication dataset in the MII data integration centers, marking a significant step toward future structured, detailed medication documentation.

Keywords: Neo4j, FHIR, IDMP, medication, medical informatics, interoperability, Medical Informatics Initiative Germany

Markus Budeus¹

Martin Boeker¹

1 Chair of Medical Informatics, Institute for AI and Informatics in Medicine, TUM Klinikum Rechts der Isar, Technical University of Munich, Munich, Germany

Zusammenfassung

Einführung: Die Verfügbarkeit von syntaktisch und semantisch interoperablen Daten eine entscheidende Voraussetzung für die Durchführung hochwertiger Forschung in der Medizin. Die Medizininformatik-Initiative stellt für Medikationsdaten mit ihrem FHIR-basierten Kerndatensatz ein effektives Datenmodell bereit, welches ein hohes Maß an Strukturiertheit und Datenqualität ermöglicht. Allerdings sind sehr detaillierte Informationen über Medikamente notwendig, um das Potenzial dieses Datenmodells auszuschöpfen. Diese liegen in den Datenintegrationszentren in der Regel nicht vor.

Zielsetzung: Um dieses Problem anzugehen, setzt sich diese Studie zum Ziel, zum Kerndatensatz konforme FHIR *Medication*-Objekte aus Medikamenten-Datenbanken zu generieren und die Konformität der generierten Daten zu gegebenen Standards zu überprüfen.

Methoden: Medikamentendaten aus dem MMI Pharmindex wurden in eine Neo4j-Graphdatenbank geladen und durch manuelle sowie automatische Mapping-Techniken mit terminologischen Codes angereichert. Die resultierenden graph-strukturierten Daten wurden dann zu FHIR-Objekten konvertiert.

Ergebnisse: Die Konvertierung lieferte rund 70.000 zum Kerndatensatz konforme FHIR *Medication*-Objekte. Die Anreicherung mit terminologischen Codes und die Einhaltung der IDMP-Standards waren nur teilweise erfolgreich, wobei unvollständige Quelldaten als Hauptproblem identifiziert wurden.

Schlussfolgerung: Unsere Studie zeigt den Bedarf nach qualitativ hochwertigeren Daten zu in Deutschland eingesetzten Medikamenten auf. Dennoch ermöglichen die generierten Daten den Datenintegrationszentren, schneller vollständige Informationen zum Einsatz von Medikamenten Kerndatensatz-konform zu erfassen, was einen wichtigen Schritt zu strukturierter und detaillierter Dokumentation von Medikamenteneinsatz darstellt.

Schlüsselwörter: Neo4j, FHIR, IDMP, Medikationsdaten, medizinische Informatik, Interoperabilität, Medizininformatik-Initiative

Introduction

High-quality medical research requires access to large amounts of syntactically and semantically interoperable data. An extensive use of national and international standards and terminologies is critical to achieve the level of interoperability required for this use case [1]. However, for medication data, achieving this level of quality is an exceptionally difficult task [2]. In many German hospitals, a significant portion of medication data is still documented in an unstructured or non-standardized form without the use of terminologies [3]. One likely reason is that clinical documentation is created for clinical, not research, purposes. But automatic data processing is far more important when handling large datasets, which is more common in research than clinical use cases. Despite these issues, several standards and terminologies already exist which can help pave the way to interoperable medication data. For example, active agents in medications can be encoded using the Global Substance Registration System (GSRS) Unique Ingredient Identifier (UNII) [4] or via the Chemical Abstracts Service (CAS) Registry Number [5]. Pharmaceutical products can be described by the German “Pharmazentralnummer” (PZN) or, on an international level, the RxNorm Concept Unique Identifier (RXCUI) [6].

On a higher level, the Identification of Medicinal Products (IDMP) standards, and HL7’s Fast Healthcare Interoperability Resources (FHIR) [7] provide a framework for the structured representation of medicinal data. On top of that, the German Medical Informatics Initiative [8] has created the “Basismodul Medikation” [2], [9] as part of its core data set (CDS), which refines the FHIR medication structure based on the IDMP standards. This CDS requires the use of terminologies, e.g. the PZN, to encode medications and must be implemented by the Data Integration Centers (DICs) that are part of the MII [8]. As an initial use case, Zabka et al. [2] and Sass et al. [10] have previously mapped “Operationen- und Prozedurenschlüssel” (OPS) codes, which are used for reimbursement purposes in Germany, to the MII CDS resources to facilitate easier adoption of structured medication documentation. Unfortunately, medication-related OPS codes only exist for a

limited number of medications, as most are irrelevant to reimbursement. Mapping data from existing clinical documentation to the CDS Medication remains a difficult task for the Data Integration Centers of the German MII. The complexity of this task could be reduced by provisioning a set of pre-defined, reusable FHIR *Medication* instances with unique, standardized, and coded identifiers for the medications used in modern healthcare. Then, the task of encoding medication information in FHIR boils down to identifying the appropriate pre-existing FHIR *Medication*.

The objective of this work is (1) to design and implement a software that generates such a set of IDMP- and MII CDS-compliant FHIR *Medication* objects based on pharmaceutical databases and (2) to evaluate the generated resources in regard to completeness, semantic conformity with standards and adherence to the MII CDS.

Methods

The requirements for the information model and the final FHIR implementation were derived from several sources. The information model needs to be compliant with the IDMP standards to the greatest possible extent. This requirement is in accordance with European law, which stipulates that the standards set out in IDMP must be adhered to for the purposes of electronic exchange of pharmaceutical data [11]. In addition, the generated FHIR objects must adhere to the 2023 MII CDS Medication Module. On a technical level, the generated FHIR resources must be easily implementable by the MII DICs into their respective ETL processes to ensure adoption of the resources.

Table 1 lists the terminologies we used in the context of this work.

The MMI Pharmindex was used as primary data source for terminology, providing the following information needed for the generation of Medication FHIR instances:

- names and PZNs of pharmaceutical products
- manufacturer information of pharmaceuticals
- ingredients of pharmaceuticals, including their dosage and ASK number, if known

- the dose form of pharmaceuticals
- ATC codes

Notably, the MMI Pharmindex has a restricted licensing model [12] that prohibits the central publication of generated FHIR instances, but requires their local generation. Therefore, an application was developed that generates FHIR resources locally using the MMI Pharmindex data as input, which can be distributed to all DICs for on-site generation of their Medication resources. Instead of directly converting the data into the FHIR format, we first load the relevant information into a graph database. We chose this approach for several reasons:

- A graph-based representation can be derived naturally from the hierarchical structures of real-world medication processes and documentation.
- The graph structure aligns well with IDMP, whose standards also describe real-world products with graphs.
- As detailed information on medication and drugs is not always available, the data in the MMI Pharmindex and other corresponding resources are often incomplete. Using a graph database, missing information could be omitted without introducing complicated handling of missing values.
- A graph representation can be easily extended and allowed the linkage to supporting information from additional data sources without modifying the schema used by existing data.
- Relational queries requiring the JOIN of multiple raw MMI Pharmindex dataset tables were much easier to handle with the highly optimized graph-based query language of Neo4J (Cypher).

We designed the graph database model based on a real-world model of pharmaceuticals, taking into account requirements imposed by the MMI Pharmindex's existing structure. We identified the FHIR resource types *Medication*, *Substance* and *Organization* as relevant top-level entities. A *Medication* represents a pharmaceutical product and contains its name, dose form, and ingredients with amounts. A *Substance* represents a pharmaceutical substance, in our context often a pharmaceutically active compound like acetylsalicylic acid, or a carrier substance

like water. An *Organization*, in our context, is the manufacturer of a specific pharmaceutical product. FHIR version R4, which is currently in use at the MII, was also used for this project.

Ambiguity in the data was reduced to a minimum by refining the data set with as many terminological identifiers as possible. We were unable to identify an automated way to include UCUM codes for units and EDQM Standard Terms for dose forms given in the MMI Pharmindex. Due to their high relevance to the IDMP standards, we manually mapped units and dose forms to UCUM and EDQM. We skipped all elements for which we could not find an appropriate mapping.

Regarding other terminologies, the MMI Pharmindex already contains PZNs, ATC codes, and ASK numbers. For all pharmaceutical substances where the ASK number is given in the MMI Pharmindex, we acquired the primary CAS number and the INN from the BfArM's AMIce "Stoffbezeichnungen Rohdaten" [13] dataset. For all substances with a primary CAS number in the AMIce dataset, we used the REST API of the public GSRS instance to find the UNII of the respective substance. If the GSRS returned multiple substances, all of which were coded using the corresponding CAS number, the following algorithm was used:

1. Substances for which the CAS number in question was not of type PRIMARY were excluded
2. If multiple substances remained, all substances with status "non-approved" were excluded
3. If multiple substances remained, all substances of class "concept" were excluded
4. If more than one substance remained, or if all substances had been excluded at any point, the match was discarded.

The data loading application was implemented using the Java programming language. Alongside the MMI Pharmindex data, we included the handcrafted mapping tables for the MMI Pharmindex units and dose forms into our application to construct the graph database. We also included a mapping table containing the substance ids from the MMI Pharmindex and the UNII, GSRS substance name, CAS numbers and RXCUIs from the GSRS. The

Table 1: Terminologies included in the graph, the type of concept they describe, and scope of national or international relevance

Terminology	Short	Describes	Scope
Pharmazentralnummer [17]	PZN	pharmaceutical product	Germany
Anatomical Therapeutic Chemical Classification System [18]	ATC	pharmaceutical product	International
EDQM Standard Terms [19]		dose form	International
Unified Code for Units of Measure [20]	UCUM	units of measurement	International
Arzneistoffkatalog-Nummer [13]	ASK	pharmaceutical substance	Germany
Chemical Abstracts Service Registry Number [5]	CAS	pharmaceutical substance	International
International Nonproprietary Name [21]	INN (German version)	pharmaceutical substance	Germany
RxNorm Concept Unique Identifier [6]	RXCUI	pharmaceutical substance	International
Unique Ingredient Identifier [4]	UNII	pharmaceutical substance	International

process of extracting data from the GSRS API and building this mapping table was executed on December 2nd, 2023 [14]. According to the licensing regulations of the AMLce “Stoffbezeichnungen Rohdaten” database, the file can be supplied as a file to our application at runtime. With that, all known CAS numbers and the INN were imported from that dataset for substances linked to the MMI data via the ASK number.

We used Neo4j as graph database due to easy interaction with the Cypher query language and its Java driver that enables programmatic access to the graph database. The Java application was extended by adding a feature that reads the graph database and constructs FHIR *Medication*, *Organization*, and *Substance* instances and writes them in JSON format to a designated location on the local file system.

To populate the graph database, we used the MMI Pharminx version 1.36 and the AMLce “Stoffbezeichnungen Rohdaten” file from 2023.

For the export, all nodes from the graph database which resemble entries in the table for pharmaceutical products of the MMI Pharminx (PRODUCT.CSV), are used as candidates for the generation of a FHIR *Medication*. Products that are incompatible with the MII CDS for any reason are skipped during the export. In the same way, we converted nodes generated from the table for molecule descriptions (MOLECULE.CSV) into FHIR *Substances* and data from the table for organizational entities (COMPANY.CSV) to *Organization* resources. We analyzed the presence of PZN and ATC codes on the generated *Medication* instances, as well as the presence of ASK numbers, CAS numbers, UNII, RXCUIs, and INNs on the generated *Substance* instances. We also checked the number of units and dose forms in the MMI Pharminx for which we could provide a UCUM or EDQM Standard Term mapping, respectively. We further analyzed how often our FHIR instances contain dose forms with an EDQM Standard Terms mapping and UCUM-encoded units. Lastly, we used simplifier validate [15] to test conformance of the generated resources to the MII CDS. Because a single validation takes multiple seconds, we only validated a limited number of randomly selected FHIR *Medications*, *Substances*, and *Organizations*. We ensured to validate at least one *Substance* instance which has no assigned terminological code, and at least one *Substance* with all supported terminological codes (ASK, CAS, UNII, RXCUI, and INN).

Results

At the core of our model of real-world pharmaceuticals is the drug, which represents a single unit of a medication. For example, this could be a tablet or an ampoule with a liquid for injection. We modeled the designated means of administration as reference to a dose form. Drugs are sold as pharmaceutical products in packages manufactured by a pharmaceutical company. A drug comprises ingredients, which are modeled as a combination of a

specific substance and amount, given as quantity and unit. A substance represents a chemical compound and has a name. Because the same units are reused for many ingredient amounts, we modeled them as a separate entity. Figure 1 displays this conceptual model. The semantics of most concepts are provided by standardized terminologies. For example, a PZN is a standardized identifier for a pharmaceutical product, ASK numbers and CAS Registry Numbers assign identifiers to pharmaceutical substances, and UCUM provides standardized codes for units. In most cases, pharmaceutical entities are unambiguously identified with corresponding codes. A major exception is the anatomical, therapeutic, and chemical “classification” (ATC) multiaxial system: By its structure, a single drug can have more than one ATC code. ATC is therefore misaligned for unambiguous identification and selection of drugs, however, has still high clinical relevance for documentation of medication.

Based on the presented real-world model, we developed a graph model and populated it with the data from the MMI Pharminx. Figure 2 presents the graph representation of a sample medication in our graph model, which employs small modifications from the real-world model for practical reasons. Unit nodes were generated based on the units given in the MMI Pharminx. Where a mapping was found, we added a corresponding UCUM code and a UCUM label to the respective nodes. The dose form mapping frequently mapped multiple MMI Pharminx dose forms to the same EDQM Standard Term, making adding the EDQM code to the respective nodes impractical. Therefore, the EDQM Standard Terms are modeled as separate nodes which are linked to respective MMI Pharminx dose forms via a relation. All remaining codes reference the concept they describe via a relation to the respective concept. The only exception is the ATC code, which is not suitable for unambiguously describing a drug for the reasons previously outlined. Consequently, it has a different relationship with the *Drug* nodes. All nodes representing terminological codes share the *Code* label, attribute, and relation to a *CodingSystem* node. This lets users query the graph for all known codes for a given concept, without needing to know which terminologies are present. This further allows adding more terminologies to the graph in the future without adapting existing queries.

Based on this graph model (see Figure 3), we developed a mapping strategy for the graph contents to the FHIR objects. The FHIR *Medication* instances were populated with medication names, dose forms, and ingredient information. Detailed information about substances were written in FHIR *Substance* instances, which were then linked by the corresponding FHIR *Medication* instances. Manufacturer information was written into an *Organization* instance that also is linked by the corresponding FHIR *Medication* instances. Figure 3 displays which of the generated CDS FHIR profiles encompass the information of each node in the graph model. The next step was to create one FHIR *Substance* instance per *Substance* node from the graph, one FHIR *Organization* per *Company*

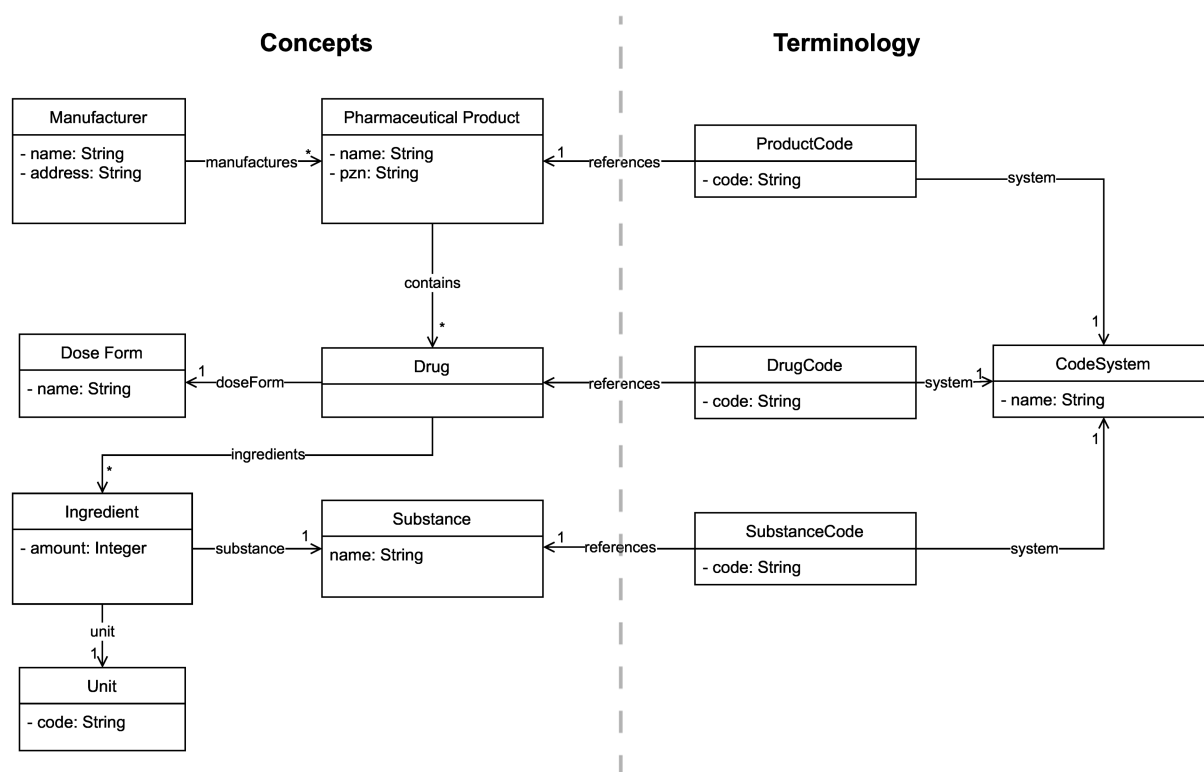


Figure 1: Simplified UML model of pharmaceutical products. Some concepts are referenced by codes from code systems like ASK, ATC or PZN. Entities in the conceptual model also correspond to similar concepts in models employed by the IDMP Standards [22] or in the organizational structure of RxNorm ([23], Fig. 3).

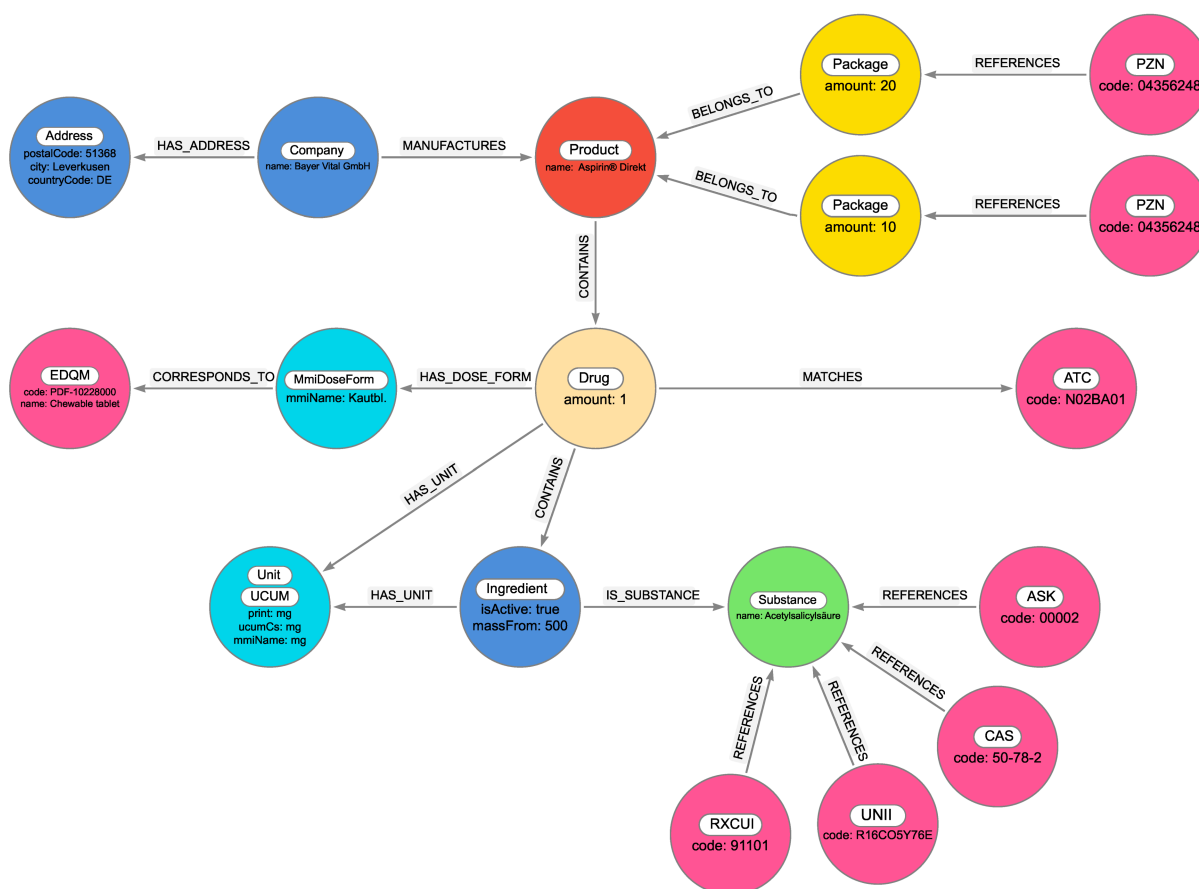


Figure 2: The pharmaceutical product “Aspirin® Direkt, 500 mg Kautablette” as represented in our graph (with minor modifications for better readability)

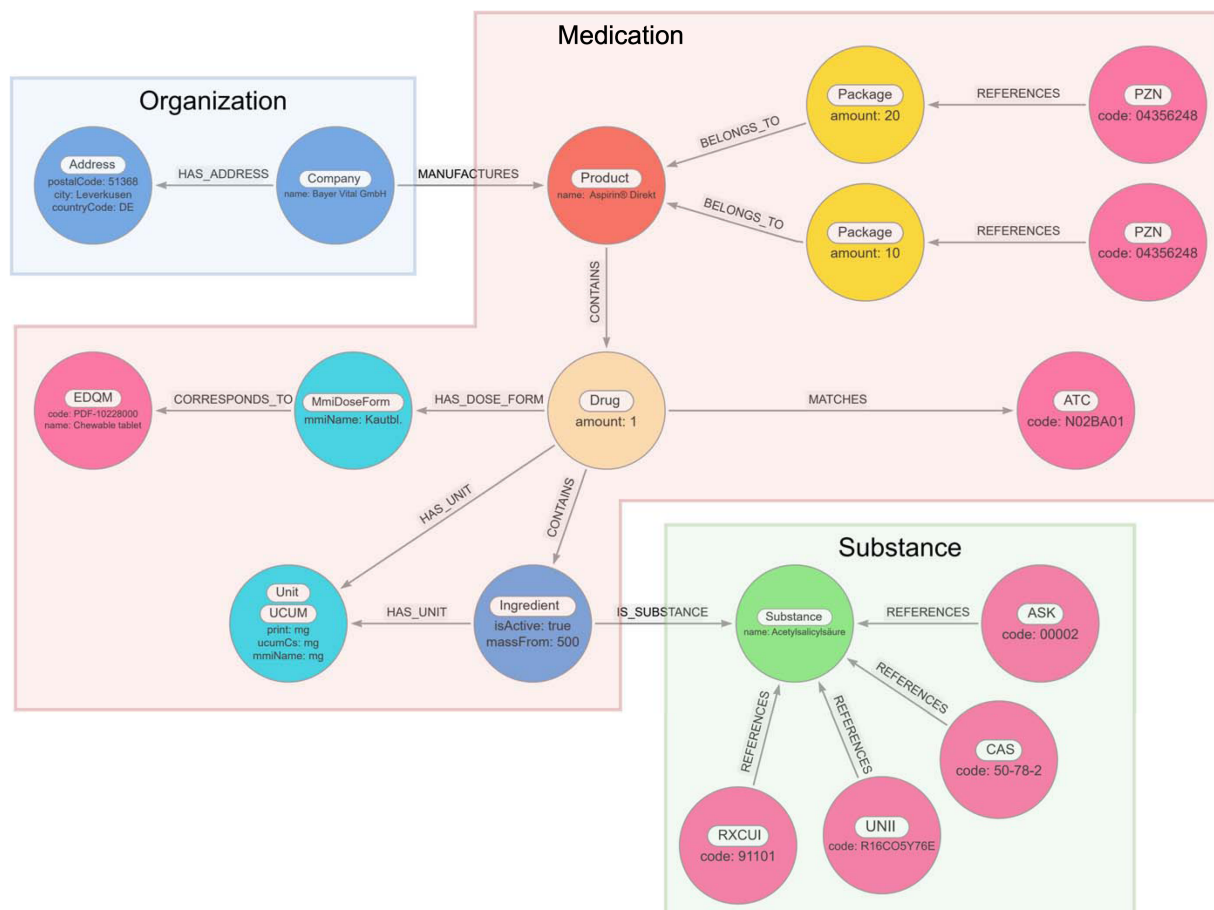


Figure 3: FHIR and IDMP conformant graph model with Product, Substance, and Company nodes and their projection onto CDS FHIR profiles Medication, Substance, and Organization

node, and one FHIR *Medication* per *Product* node. However, if a *Product* node is linked to multiple *Drug* nodes in the graph, this approach is not feasible, because each CDS FHIR *Medication* instance can only have *one* dose form and *one* set of ingredients. In order to support such cases, the MII CDS *Medication* profile allows that ingredients of a *Medication* may be linked to FHIR *Medication* instances themselves. We thus modelled such cases by creating one FHIR *Medication* instance per *Drug* node in the graph, which we refer to as composite children. Then, we add another *Medication* instance that contains information from the MMI Pharminx *Product* node, which we call a composite parent. The composite parent then links to the respective composite children. We also encountered drugs with no specified ingredients in the MMI Pharminx. As the MII CDS *Medication* profile includes the mandatory requirement to specify at least one ingredient, a conversion of these products was not possible. The export routine therefore skips them. Another issue was that we could not map all dose forms to EDQM standard terms. Reasons for that included terms being ambiguous or too vague (e.g. “Inf.-Fl.”), terms not describing a dose form (e.g. “Btl.”, probably an abbreviation for bag), or that we weren’t able to identify a suitable EDQM Standard Term (e.g. for “Pulv. z. Herst. e. Aerosols” – powder for creation of an aerosol). Out of the 386 terms for dose form in the MMI Pharminx, 302 (78.2%) received an

EDQM mapping. Similarly, we were not able to map all units found in the MMI Pharminx to a UCUM code, mapping 47 out of 75 units (62.7%) to a UCUM code. In the remaining cases, the units were either not precise enough (e.g. “measuring spoon”) or could not be identified with sufficient certainty (e.g. “E”, which is likely a shorthand for units, which is not descriptive enough to identify the corresponding UCUM unit). Because UCUM units and EDQM Standard terms are required to achieve full IDMP compliance, but the mappings are incomplete, not all generated FHIR instances are entirely IDMP compliant in this regard. The entire process generated 74191 *Medication* objects, 15901 *Substances*, and 4971 *Organizations*. Table 2 shows the number of generated FHIR *Medications* containing a dose form for which a corresponding EDQM Standard Term was found and the number of *Medications* containing only units with a UCUM code. Because multiple units can be given per FHIR *Medication* instance (one per ingredient), the table also includes the number of units mentioned in a resource and what part of these units have a UCUM code.

Another requirement for IDMP compliance is that pharmaceutical products and substances must be uniquely identifiable, which can be achieved using corresponding terminologies. Table 3 lists the number of codes from terminologies in the generated FHIR *Medication* and *Substance* objects. Due to assignments of multiple codes

Table 2: The presence of UCUM-compliant units and EDQM-compliant dose forms in generated Medication objects
(also presented in [14])

	No.	Total no.	Percentage
EDQM Standard Terms out of all dose forms	57,054	72,244	79.0%
UCUM-compliant units out of all units mentioned in <i>Medication</i> objects	158,408	190,111	83.3%
<i>Medication</i> objects containing only units with a UCUM code	43,946	74,191	59.2%

Table 3: The number of terminological identifiers present in the exported FHIR objects.
The “Total occurrences” column counts how often codes appear in the FHIR instances, whereas the “Present in” column counts a terminological code at most once per resource.

Code	Objects type	No.	Present in		Total occurrences
			No.	%	
PZN	Medication	68,587	68,586	100.0%	109,458
ATC	Medication	72,244	72,152	99.9%	72,241
ASK	Substance	15,901	6,361	40.0%	6,361
CAS	Substance	15,901	4,740	29.8%	7,354
INN	Substance	15,901	3,230	20.3%	4,319
RXCUI	Substance	15,901	3,756	23.6%	4,051
UNII	Substance	15,901	4,319	27.2%	3,230

of the same terminology to the same concept, the total number of occurrences of terminology codes can be higher than the number of FHIR objects.

Almost every *Medication* is identified via a corresponding PZN. For the *Substances*, only 40% of instances contain an ASK number. Due to the strategy to resolve the remaining terminological identifiers relying on the ASK number, the remaining 60% of substances have no known terminological identifier. For the FHIR validation via the Simplifier validator [15], we randomly selected 30 Medication objects that are neither composite parents nor composite children. We further validated 20 composite parent Medications, composite children, *Organization* instances and *Substance* instances each. All passed the validation, leading us to assume all generated FHIR resources adhere to the MII CDS Medication module with sufficient certainty.

Discussion

We have created a software that can enrich data from the MMI Pharminx with information from the AMIce database, the GSRs, UCUM codes, and EDQM standard terms to transform this data into corresponding MII CDS FHIR *Medication* objects. Despite the need to filter out many pharmaceutical products incompatible with the MII CDS and shortcomings regarding IDMP compliance, we created FHIR *Medication* instances covering a large portion of the pharmaceuticals used in Germany.

This is the first time detailed information on medication was provided as standardized MII CDS FHIR *Medication* objects with this degree of coverage. Although previous works [2], [10] have presented a way to generate FHIR *Medication* instances from OPS codes, our approach is no longer limited to medications for which OPS codes exist.

This can help the MII DICs provide structured medication-related documentation in the FHIR format, as they no longer need to collect all information about pharmaceutical products individually, but instead only need to identify the medicinal products, e.g. via PZN, used at the individual hospitals. We expect the issue of incomplete data to only moderately affect the practical usability of the dataset. For one, preliminary observations suggest that common medications tend to have more complete source data and thus more often result in IDMP-compliant FHIR *Medications*. We therefore expect the frequency at which missing or incomplete *Medication* resources are needed to be lower than the nominal occurrence frequency we measured. When it happens regardless, DICs can still use the non-compliant resources or generate the required *Medication* on-demand with whatever data is available. Future work can focus on improving available source data on medications to further alleviate this issue.

Still, our results have notable limitations. Although we evaluated the completeness of data elements to assess the IDMP compliance and usability of our results, other metrics were not considered. Any errors in the data sources we used remain present in the generated FHIR objects. In addition, our results are only usable by parties with access to the MMI Pharminx. As all DICs have access to the MMI Pharminx via the MII, this is not currently an issue. It also must be explored to what extent the *Medications* may be shared with researchers, given they contain data from the MMI Pharminx. Next, no mechanism exists as of now to keep the medications up to date with the constantly changing market. Users of our software can regularly recreate the graph using the newest MMI Pharminx dataset, but this will not reflect changes in the referenced terminologies. The manual mappings of dose forms and units to UCUM codes and EDQM Standard Terms are prone to errors and have not

been subjected to validation. Further research may focus on these mapping tables to improve their completeness and evaluate the correctness of mappings. In addition, other ways of linking pharmaceutical substances and products to terminology codes can be explored to improve the availability of these codes in the generated FHIR *Medication* instances. Further work also includes validating the applicability of the *Medications* to a range of different real-world settings.

Conclusion

We created a software capable of transforming data from the MMI Pharminindex and other databases into MII CDS-compliant FHIR *Medication* instances. Although the data on many pharmaceuticals were found to be incompatible with the MII CDS and others lacked the necessary information for IDMP compliance, the created MII CDS FHIR *Medication* instances cover a wide range of pharmaceuticals on the German market. This allows the MII DICs to use predefined FHIR objects for medication documentation and can help accelerate the provision of a complete dataset on medication in MII and NUM. Our goal of delivering pre-defined, MII CDS- and IDMP-compliant FHIR *Medications* has been fulfilled with respect to incomplete source data. This marks a significant step toward future structured, detailed medication documentation in the MII and NUM that will enable adequate retrospective multisite studies.

Notes

Code availability

A version of the application described in this paper is publicly available [16].

Funding

This study was funded by the German Federal Ministry of Research, Technology and Space (BMFTR) within the “Medical Informatics Funding Scheme” (FKZ 01ZZ2304A).

Competing interests

The authors declare that they have no competing interests.

References

1. Pimenta N, Chaves A, Sousa R, Abelha A, Peixoto H. Interoperability of Clinical Data through FHIR: A review. *Procedia Comput Sci.* 2023 Jan 1;220:856-61.
2. Zabka S, Ammon D, Ganslandt T, Gewehr J, Haverkamp C, Kiefer S, Lautenbacher H, Löbe M, Thun S, Boeker M. Towards a Medication Core Data Set for the Medical Informatics Initiative (MII): Initial Mapping Experience between the German Procedure Classification (OPS) and the Identification of Medicinal Products (IDMP). In: *Proceedings of the Joint Ontology Workshops 2019 (JOWO 2019)*; 2019 Sep 23-25; Graz, Austria. (CEUR Workshop Proceedings; 2518).
3. Vass A, Reinecke I, Boeker M, Prokosch HU, Gulden C. Availability of Structured Data Elements in Electronic Health Records for Supporting Patient Recruitment in Clinical Trials. In: Otero P, Scott P, Martin SZ, Huesing E, editors. *MEDINFO 2021: One World, One Health – Global Partnership for Digital Innovation*. IOS Press; 2022. (Studies in Health Technology and Informatics; 290). p. 130-4. DOI: 10.3233/SHTI220046
4. Peryea T, Southall N, Miller M, Katzel D, Anderson N, Neyra J, Stemann S, Nguyen DT, Amugoda D, Newatia A, Ghazzaoui R, Johanson E, Diederik H, Callahan L, Switzer F. Global Substance Registration System: consistent scientific descriptions for substances related to health. *Nucleic Acids Res.* 2021 Jan 8;49(D1):D1179-D1185. DOI: 10.1093/nar/gkaa962
5. Hamill KA, Nelson RD, Vander Stouw GG, Stobaugh RE. Chemical Abstracts Service Chemical Registry System. 10. Registration of substances from pre-1965 indexes of Chemical Abstracts. *J Chem Inf Comput Sci.* 1988 Nov;28(4):175-9. DOI: 10.1021/ci00060a002
6. Nelson SJ, Zeng K, Kilbourne J, Powell T, Moore R. Normalized names for clinical drugs: RxNorm at 6 years. *J Am Med Inform Assoc.* 2011 Jul-Aug;18(4):441-8. DOI: 10.1136/amiajnl-2011-000116
7. HL7 International. Index – FHIR v4.0.1. HL7 International; [updated 2019 Nov 1, cited 2025 Mar 28]. Available from: <https://hl7.org/fhir/R4/index.html>
8. Semler SC, Wissing F, Heyder R. German Medical Informatics Initiative. *Methods Inf Med.* 2018 Jul;57(S 01):e50-e56. DOI: 10.3414/ME18-03-0003
9. Boeker M, Thun S, Saß J, Zabka S, Zautke A. MII – Basismodul Medikation (2023). SIMPLIFIER.NET; 2023 [cited 2025 Mar 28]. Available from: <https://simplifier.net/medizininformatikinitiative-modulmedikation>
10. Sass J, Zabka S, Essenwanger A, Schepers J, Boeker M, Thun S. Fast Healthcare Interoperability Resources (FHIR) Representation of Medication Data Derived from German Procedure Classification Codes (OPS) Using Identification of Medicinal Products (IDMP) Compliant Terminology. In: Röhrig R, Reißbarth T, Brannath W, Prokosch HU, Schmidtman I, Stolpe S, Zapf A, editors. *German Medical Data Sciences: Bringing Data to Life. Proceedings of the Joint Annual Meeting of the German Association of Medical Informatics, Biometry and Epidemiology (gmds e.V.) and the Central European Network – International Biometric Society (CEN-IBS) 2020 in Berlin, Germany*. IOS Press; 2021. (Studies in Health Technology and Informatics; 278). p. 231-6. DOI: 10.3233/SHTI210074
11. Regulation (EU) No 520/2012. Art. 25, 26.
12. Vidal MMI Germany GmbH. MMI Pharminindex Rohdaten. Langen: Vidal MMI; [cited 2025 Mar 28]. Available from: <https://www.mmi.de/mmi-pharminindex/mmi-pharminindex-daten>

13. Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM). Arzneimittelinformationssystem AMLce. Bonn: BfArM; [updated 2025; cited 2025 Apr 01]. Available from: https://www.bfarm.de/DE/Arzneimittel/Arzneimittelinformationen/Arzneimittel-recherchieren/AMLce/_node.html
14. Budeus M. Matching of real-world medication data based on a graph-based representation of an IDMP conformant medication model [unpublished master's thesis]. Augsburg: University of Augsburg; 2024.
15. SIMPLIFIER.NET. [updated 2025; cited 2025 Apr 01]. Available from: <https://simplifier.net/validate>
16. Budeus M. medizininformatik-initiative/Medication-Graph-FHIR-Converter. Version 1.0.2. Zenodo; 2025. DOI: 10.5281/zenodo.15746125
17. IFA GmbH. IFA-Informationen zur PZN. IFA GmbH; [cited 2025 Apr 01]. Available from: <https://www.ifaffm.de/de/ifa-fuer-anbieter/ifa-informationen.html>
18. World Health Organization. Anatomical Therapeutic Chemical (ATC) Classification. World Health Organization; [cited 2025 Apr 01]. Available from: <https://www.who.int/tools/atc-ddd-toolkit/atc-classification>
19. Council of Europe – European Directorate for the Quality of Medicines & HealthCare. Standard Terms Database. Strasbourg: Council of Europe; cited 2025 Apr 01]. Available from: <https://www.edqm.eu/en/standard-terms-database>
20. Schadow G, McDonald CJ. The unified code for units of measure. Indianapolis, IN, USA: Regenstrief Institute and UCUM Organization; 2009. p. 99.
21. Kopp-Kubel S. International Nonproprietary Names (INN) for pharmaceutical substances. Bull World Health Organ. 1995;73(3):275-9.
22. Tschorn U. IDMP Definition – IDMP Wiki. IDMP1 GmbH; [updated 2015 May 22, cited 2025 Apr 01]. Available from: <https://www.idmp1.com/wiki/idmp-definition/>
23. Bodenreider O, Cornet R, Vreeman DJ. Recent Developments in Clinical Terminologies – SNOMED CT, LOINC, and RxNorm. Yearb Med Inform. 2018 Aug;27(1):129-39. DOI: 10.1055/s-0038-1667077

Corresponding author:

Markus Budeus

Chair of Medical Informatics, Institute for AI and Informatics in Medicine, TUM Klinikum Rechts der Isar, Ismaninger Straße 22, 81675 Munich, Germany
markus.budeus@tum.de

Please cite as

Budeus M, Boeker M. Graph-based creation of IDMP-conformant FHIR Medication instances using Neo4j: An evaluation study. *GMS Med Inform Biom Epidemiol.* 2025;21:Doc10. DOI: 10.3205/mibe000282, URN: urn:nbn:de:0183-mibe000282

This article is freely available from

<https://doi.org/10.3205/mibe000282>

Published: 2025-09-09

Copyright

©2025 Budeus et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 License. See license information at <http://creativecommons.org/licenses/by/4.0/>.